

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 13:15:50 ; Search time 172 Seconds
(without alignments)
2188.245 Million cell updates/sec

Title: US-10-038-972A-13

Perfect score: 3994

Sequence: 1 MAADGYLPDWMLEDTLSEGR.....TNGVSEPRPIGTRYLRNL 735

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3994	100.0	735	2	O56652	O56652 adeno-ssoc
2	3973	99.5	735	2	Q6JC41	Q6JC41 adeno-ssoc
3	3972	99.4	735	2	Q6JC06	Q6JC06 adeno-ssoc
4	3969	99.4	735	2	Q6JC20	Q6JC20 adeno-ssoc
5	3968	99.3	735	2	Q6JC09	Q6JC09 adeno-ssoc
6	3967	99.3	735	2	Q6JC11	Q6JC11 adeno-ssoc
7	3961	99.2	735	2	Q6JC01	Q6JC01 adeno-ssoc
8	3961	99.2	735	2	Q6JC21	Q6JC21 adeno-ssoc
9	3955	99.0	735	2	Q6JC25	Q6JC25 adeno-ssoc
10	3954	99.0	735	2	Q6JC25	Q6JC25 adeno-ssoc
11	3953	99.0	735	2	Q6JB29	Q6JB29 adeno-ssoc
12	3946	98.8	735	2	Q670R4	Q670R4 adeno-ssoc
13	3945	98.8	735	2	Q6JC07	Q6JC07 adeno-ssoc
14	3941	98.7	735	2	Q670Q6	Q670Q6 adeno-ssoc
15	3940	98.6	735	2	Q6JC25	Q6JC25 adeno-ssoc
16	3939	98.6	735	2	Q6JC26	Q6JC26 adeno-ssoc
17	3937	98.6	735	2	Q670R6	Q670R6 adeno-ssoc
18	3934	98.5	735	2	Q670S0	Q670S0 adeno-ssoc
19	3933	98.5	735	2	Q6JC29	Q6JC29 adeno-ssoc
20	3932	98.4	735	2	Q6JB24	Q6JB24 adeno-ssoc
21	3931	98.4	735	2	Q670R8	Q670R8 adeno-ssoc
22	3928.5	98.4	734	2	Q6JC00	Q6JC00 adeno-ssoc
23	3927	98.3	735	2	Q670R0	Q670R0 adeno-ssoc
24	3923	98.2	735	2	Q6JC35	Q6JC35 adeno-ssoc
25	3921	98.2	735	2	Q670Q7	Q670Q7 adeno-ssoc
26	3917	98.1	735	2	Q6JC31	Q6JC31 adeno-ssoc
27	3916	98.0	735	2	Q6JC27	Q6JC27 adeno-ssoc
28	3916	98.0	735	2	Q6JC30	Q6JC30 adeno-ssoc
29	3904	97.7	735	2	Q6JC33	Q6JC33 adeno-ssoc
30	3725	93.3	735	2	Q6JB27	Q6JB27 adeno-ssoc
31	3711	92.9	735	2	Q6JC34	Q6JC34 adeno-ssoc

32	3709.5	92.9	734	2	Q6JC02	Q6JC02 adeno-ssoc
33	3706	92.8	735	2	Q6JC17	Q6JC17 adeno-ssoc
34	3704.5	92.8	734	2	Q6JC04	Q6JC04 adeno-ssoc
35	3704	92.7	735	2	Q6JB21	Q6JB21 adeno-ssoc
36	3704	92.7	735	2	Q6JC42	Q6JC42 adeno-ssoc
37	3703	92.7	735	2	Q6JB26	Q6JB26 adeno-ssoc
38	3701	92.7	735	2	Q6JC44	Q6JC44 adeno-ssoc
39	3700	92.6	735	2	Q6JC28	Q6JC28 adeno-ssoc
40	3690.5	92.4	734	2	Q6JC03	Q6JC03 adeno-ssoc
41	3688	92.3	735	2	Q6JC38	Q6JC38 adeno-ssoc
42	3685	92.3	735	2	Q6JC36	Q6JC36 adeno-ssoc
43	3683	92.2	735	2	Q670Q8	Q670Q8 adeno-ssoc
44	3683	92.2	735	2	Q6JC43	Q6JC43 adeno-ssoc
45	3676	92.0	735	2	Q6JB20	Q6JB20 adeno-ssoc

ALIGNMENTS

RESULT 1
O56652 PRELIMINARY; PRT; 735 AA.
AC O56652;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Major coat protein VP1.
OS Adeno-associated virus 2 (AAV2).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_taxid=10804;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95088582; PubMed=7996133;
RA Ruffing M., Heid H., Kleinschmidt J.A.;
RT "Mutations in the carboxy terminus of adeno-associated virus 2 capsid proteins affect viral infectivity: lack of an RGD integrin-binding motif."
RT J. Gen. Virol. 75:0-0(0).
RL [2]
RN SEQUENCE FROM N.A.
RP
RA Berns K.I., Bohenzky R.A., Cassinotti P., Colvin D., Donahue B.A.,
RA Dull T., Horer M., Kleinschmidt J.A., Ruffing M., Snyder R.O.,
RA Tratschin J.-D., Weitz M.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF043303; AAC03780.1; -;
DR GO; GO:0019928; C: viral capsid; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
KW Coat protein.
SQ SEQUENCE 735 AA; 81944 MW; 980BEEF46908390B CRC64;

Query Match 100.0%; Score 3994; DB 2; Length 735;
Best Local Similarity 100.0%; Pred. No. 1.4e-242;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAADGYLPDWMLEDTLSEGRQWKLKPGPPPKPAERHKDSDRGVLVPGYKVLGPFENGLD	60
Db	1	MAADGYLPDWMLEDTLSEGRQWKLKPGPPPKPAERHKDSDRGVLVPGYKVLGPFENGLD	60
Qy	61	KGEVNEADAALAEHDKAYDRQLSDGNPNLYKYNHADAERLKEKEDTSFGNGLGRAVFQ	120
Db	61	KGEVNEADAALAEHDKAYDRQLSDGNPNLYKYNHADAERLKEKEDTSFGNGLGRAVFQ	120
Qy	121	AKKRVLPGLVPEPVKTAPKKPVEHSPVPESSSGTGKAGQOPARKLNFGQTGDAD	180
Db	121	AKKRVLPGLVPEPVKTAPKKPVEHSPVPESSSGTGKAGQOPARKLNFGQTGDAD	180
Qy	181	SVDPDQPLGGPPAPSPGLGTNTMTATGSGAPMADNNEGADVGNSSGNHCDSTWMDRVI	240
Db	181	SVDPDQPLGGPPAPSPGLGTNTMTATGSGAPMADNNEGADVGNSSGNHCDSTWMDRVI	240
Qy	241	TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPWGYFDNFRFCHFSPRDWQRLI	300

```
Db 241 TTSTRWALPTNNHLYKQISSQSGASNDNHFGYSTPWGYDFNRFCHFSRDMQRLI 300
QY 301 NNNWGRPKRLNFKLNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQG 360
Db 301 NNNWGRPKRLNFKLNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQG 360
QY 361 CLPPPPADVFVMPQYGYLTNNQSGAVGRSSFYCLEYFSSQMLRTGNNFTFSTYTFEDVVF 420
Db 361 CLPPPPADVFVMPQYGYLTNNQSGAVGRSSFYCLEYFSSQMLRTGNNFTFSTYTFEDVVF 420
QY 421 HSSYAHQSLSRLMNLPLIDQYLYLSRTNTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPG 480
Db 421 HSSYAHQSLSRLMNLPLIDQYLYLSRTNTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPG 480
QY 481 PCYRQORVSKTSADNNNSYSWTGATKYHLNGRDSLVPNGPAMASHKDDDEKFFPQSGVYL 540
Db 481 PCYRQORVSKTSADNNNSYSWTGATKYHLNGRDSLVPNGPAMASHKDDDEKFFPQSGVYL 540
QY 541 IFKGQSEKTNVDIEKWMITDEEIRITNPVATEQYGSVSTNLQRGNRAATADVNTQGV 600
Db 541 IFKGQSEKTNVDIEKWMITDEEIRITNPVATEQYGSVSTNLQRGNRAATADVNTQGV 600
QY 601 LFGMWQDRDVLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
Db 601 LFGMWQDRDVLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
QY 661 FSAAKFASFTQYSTQGVSVIEWELOKENSXKNWPEIQYTSNYSKSVNVDFVTDTNGVY 720
Db 661 FSAAKFASFTQYSTQGVSVIEWELOKENSXKNWPEIQYTSNYSKSVNVDFVTDTNGVY 720
QY 721 SEPRPIGTRYLTRNL 735
Db 721 SEPRPIGTRYLTRNL 735

RESULT 2
Q6JC41 PRELIMINARY; PRT; 735 AA.
ID Q6JC41
AC Q6JC06;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE Capsid protein VPI.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388(2004).
DR EMBL; AY530578; AAS99263.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 735 AA; 81748 MW; 1555801B683FE750 CRC64;

Query Match 99.58; Score 3973; DB 2; Length 735;
Best Local Similarity 99.58; Pred. No. 3e-241;
Matches 731; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAADGYLPDWLEDTLSGIRQWKKLKEGPPPPKPAERHKDDSRGLVLPQYKYLGPFGNGLD 60
Db 1 MAADGYLPDWLEDTLSGIRQWKKLKEGPPPPKPAERHKDDSRGLVLPQYKYLGPFGNGLD 60
QY 61 KGEPVNEADAAAEHDKAYDRQLDSGNDPNLYKYNHADAEEQERLKEDTSFGNLRGAVFQ 120
```

```
Db 61 KGEPVNEADAAAEHDKAYDRQLDSGNDPNLYKYNHADAEEQERLKEDTSFGNLRGAVFQ 120
QY 121 AKKRVLLEPLGLVEEPVKTPACKKRPVHSPVEPSSSCTGKAGQQPARKRLNFGGTGDAD 180
Db 121 AKKRVLLEPLGLVEEPVKTPACKKRPVHSPVEPSSSCTGKAGQQPARKRLNFGGTGDAD 180
QY 181 SVPDPQPLGQCPAAPSGLGTNTMATGSGAPMADNNEGADGVGNSNGNHCHDSTWMDRVI 240
Db 181 SVPDPQPLGQCPAAPSGLGTNTMATGSGAPMADNNEGADGVGNSNGNHCHDSTWMDRVI 240
QY 241 TTSTRWALPTNNHLYKQISSQSGASNDNHFGYSTPWGYDFNRFCHFSRDMQRLI 300
Db 241 TTSTRWALPTNNHLYKQISSQSGASNDNHFGYSTPWGYDFNRFCHFSRDMQRLI 300
QY 301 NNNWGRPKRLNFKLNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQG 360
Db 301 NNNWGRPKRLNFKLNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQG 360
QY 361 CLPPPPADVFVMPQYGYLTNNQSGAVGRSSFYCLEYFSSQMLRTGNNFTFSTYTFEDVVF 420
Db 361 CLPPPPADVFVMPQYGYLTNNQSGAVGRSSFYCLEYFSSQMLRTGNNFTFSTYTFEDVVF 420
QY 421 HSSYAHQSLSRLMNLPLIDQYLYLSRTNTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPG 480
Db 421 HSSYAHQSLSRLMNLPLIDQYLYLSRTNTPSGTTTQSRQLQFSQAGASDIRDQSRNWLPG 480
QY 481 PCYRQORVSKTSADNNNSYSWTGATKYHLNGRDSLVPNGPAMASHKDDDEKFFPQSGVYL 540
Db 481 PCYRQORVSKTSADNNNSYSWTGATKYHLNGRDSLVPNGPAMASHKDDDEKFFPQSGVYL 540
QY 541 IFKGQSEKTNVDIEKWMITDEEIRITNPVATEQYGSVSTNLQRGNRAATADVNTQGV 600
Db 541 IFKGQSEKTNVDIEKWMITDEEIRITNPVATEQYGSVSTNLQRGNRAATADVNTQGV 600
QY 601 LFGMWQDRDVLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
Db 601 LFGMWQDRDVLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
QY 661 FSAAKFASFTQYSTQGVSVIEWELOKENSXKNWPEIQYTSNYSKSVNVDFVTDTNGVY 720
Db 661 FSAAKFASFTQYSTQGVSVIEWELOKENSXKNWPEIQYTSNYSKSVNVDFVTDTNGVY 720
QY 721 SEPRPIGTRYLTRNL 735
Db 721 SEPRPIGTRYLTRNL 735

RESULT 3
Q6JC06 PRELIMINARY; PRT; 735 AA.
ID Q6JC06;
AC Q6JC06;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE Capsid protein VPI.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388(2004).
DR EMBL; AY530613; AAS99298.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 735 AA; 81786 MW; 802F42CBFCFEFIC1 CRC64;
```


05-JUL-2004 (Tremblrel. 27, Created)
05-JUL-2004 (Tremblrel. 27, Last sequence update)
05-JUL-2004 (Tremblrel. 27, Last annotation update)
Capsid protein VPI.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]_SEQUENCE FROM N.A.
RP PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RX Gao G., Vandenbergh L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues";
RL J. Virol. 78:6381-6388 (2004).
DR EMBL; AY530610; AAS99295.1; -;
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo coat.
DR Pfam; PF00740; Parvo coat_1.
SQ SEQUENCE 735 AA; 81817 MW; 496DF18E95CABB13 CRC64;
Query Match 99.3%; Score 3968; DB 2; Length 735;
Best Local Similarity 99.3%; Pred. No. 6.1e-241;
Matches 730; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 MAADGYPDLWEDTLSEGIROWKLGPGPPPKPAERHKDDSGVLPGYKYLPGPENGLD 60
DB 1 MAADGYPDLWEDTLSEGIROWKLGPGPPPKPAERHKDDSGVLPGYKYLPGPENGLD 60
QY 61 KGEPVNEADAAALEHDKAYDRQLDSDGNPILKYNHADAEFOERLKEDTSFGNGLGRAVFQ 120
DB 61 KGEPVNEADAAALEHDKAYDRQLDSDGNPILKYNHADAEFOERLKEDTSFGNGLGRAVFQ 120
QY 121 AKKRVLEPLGLVEEPVKTAPGKRPEVHSPVEPDSSSGTGKAGQOPARKLNFQGTGDAD 180
DB 121 AKKRVLEPLGLVEEPVKTAPGKRPEVHSPVEPDSSSGTGKAGQOPARKLNFQGTGDAD 180
QY 181 SVDPDQPLGPPAAPSGLTNTMATSGAPMADNNEGADGVNNGHCDSTWMDRVI 240
DB 181 SVDPDQPLGPPAAPSGLTNTMATSGAPMADNNEGADGVNNGHCDSTWMDRVI 240
QY 241 TTSTRTWALPTYNHLYKQISSQSGASNDHYFGYSTPMGYFDNRFHCHFSPRDWRLLI 300
DB 241 TTSTRTWALPTYNHLYKQISSQSGASNDHYFGYSTPMGYFDNRFHCHFSPRDWRLLI 300
QY 301 NNNWGFRPKRLNPKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQG 360
DB 301 NNNWGFRPKRLNPKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQG 360
QY 361 CLPPPPADVFMVPOYGYLTINNGSQAVGRSSFYCLEYFPQSOMLRTGNNTFSTPEDVFP 420
DB 361 CLPPPPADVFMVPOYGYLTINNGSQAVGRSSFYCLEYFPQSOMLRTGNNTFSTPEDVFP 420
QY 421 HSSYAHQSLSLRLMPLIDQYLYLSTNTPSGTTTQSRLOFSQAGASDIRQSRNWLPG 480
DB 421 HSSYAHQSLSLRLMPLIDQYLYLSTNTPSGTTTQSRLOFSQAGASDIRQSRNWLPG 480
QY 481 PCYRQORVSKTSADNNNSYSWTGATKYHLNGRDSLNVNPGPAMASHKDDKFFPQSGVL 540
DB 481 PCYRQORVSKTSADNNNSYSWTGATKYHLNGRDSLNVNPGPAMASHKDDKFFPQSGVL 540
QY 541 IFKGQSEKNTVDIEKVMITDEEIEITNPVATEOYGSVSTNLQRNQAATADVNTQGV 600
DB 541 IFKGQSEKNTVDIEKVMITDEEIEITNPVATEOYGSVSTNLQRNQAATADVNTQGV 600
QY 601 LPMQWQDRDVLQGPIMAKIPHTDGHFHSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
DB 601 LPMQWQDRDVLQGPIMAKIPHTDGHFHSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
QY 661 FSAAKFASPTTQYSTQVSVEIWELOKENSKRWNPEIQYTSYNKSVNVDFVTDTNGVY 720
DB 661 FSAAKFASPTTQYSTQVSVEIWELOKENSKRWNPEIQYTSYNKSVNVDFVTDTNGVY 720

661 FSAAKFASPTTQYSTQVSVEIWELOKENSKRWNPEIQYTSYNKSVNVDFVTDTNGVY 720
QY 721 SEPRPIGTRYLTRNL 735
DB 721 SEPRPIGTRYLTRNL 735
RESULT 6
Q6JC11 PRELIMINARY; PRT; 735 AA.
ID Q6JC11
AC Q6JC11;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE Capsid protein VPI.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]_SEQUENCE FROM N.A.
RP PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RX Gao G., Vandenbergh L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues";
RL J. Virol. 78:6381-6388 (2004).
DR EMBL; AY530608; AAS99293.1; -;
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo coat.
DR Pfam; PF00740; Parvo coat_1.
SQ SEQUENCE 735 AA; 81797 MW; 7A6CC0B6C0C78F0D CRC64;
Query Match 99.3%; Score 3967; DB 2; Length 735;
Best Local Similarity 99.3%; Pred. No. 7e-241;
Matches 730; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 MAADGYPDLWEDTLSEGIROWKLGPGPPPKPAERHKDDSGVLPGYKYLPGPENGLD 60
DB 1 MAADGYPDLWEDTLSEGIROWKLGPGPPPKPAERHKDDSGVLPGYKYLPGPENGLD 60
QY 61 KGEPVNEADAAALEHDKAYDRQLDSDGNPILKYNHADAEFOERLKEDTSFGNGLGRAVFQ 120
DB 61 KGEPVNEADAAALEHDKAYDRQLDSDGNPILKYNHADAEFOERLKEDTSFGNGLGRAVFQ 120
QY 121 AKKRVLEPLGLVEEPVKTAPGKRPEVHSPVEPDSSSGTGKAGQOPARKLNFQGTGDAD 180
DB 121 AKKRVLEPLGLVEEPVKTAPGKRPEVHSPVEPDSSSGTGKAGQOPARKLNFQGTGDAD 180
QY 181 SVDPDQPLGPPAAPSGLTNTMATSGAPMADNNEGADGVNNGHCDSTWMDRVI 240
DB 181 SVDPDQPLGPPAAPSGLTNTMATSGAPMADNNEGADGVNNGHCDSTWMDRVI 240
QY 241 TTSTRTWALPTYNHLYKQISSQSGASNDHYFGYSTPMGYFDNRFHCHFSPRDWRLLI 300
DB 241 TTSTRTWALPTYNHLYKQISSQSGASNDHYFGYSTPMGYFDNRFHCHFSPRDWRLLI 300
QY 301 NNNWGFRPKRLNPKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQG 360
DB 301 NNNWGFRPKRLNPKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQG 360
QY 361 CLPPPPADVFMVPOYGYLTINNGSQAVGRSSFYCLEYFPQSOMLRTGNNTFSTPEDVFP 420
DB 361 CLPPPPADVFMVPOYGYLTINNGSQAVGRSSFYCLEYFPQSOMLRTGNNTFSTPEDVFP 420
QY 421 HSSYAHQSLSLRLMPLIDQYLYLSTNTPSGTTTQSRLOFSQAGASDIRQSRNWLPG 480
DB 421 HSSYAHQSLSLRLMPLIDQYLYLSTNTPSGTTTQSRLOFSQAGASDIRQSRNWLPG 480
QY 481 PCYRQORVSKTSADNNNSYSWTGATKYHLNGRDSLNVNPGPAMASHKDDKFFPQSGVL 540
DB 481 PCYRQORVSKTSADNNNSYSWTGATKYHLNGRDSLNVNPGPAMASHKDDKFFPQSGVL 540


```
Qy 541 IFGQSGSEKTNVDIEKVMITDEEIRITNPVATEQYGSVSTNLQRGNRQAATADVNTQGV 600
Db 541 IFGQSGSEKTNVDIEKVMITDEEIRITNPVATEQYGSVSTNLQRGNRQAATADVNTQGV 600
Qy 601 LPGVMWODRDVYLOGPIWAKIPHDTGHHFSPLMGGFGLKHPPPQIILIKNTPVPANPSTT 660
Db 601 LPGVMWODRDVYLOGPIWAKIPHDTGHHFSPLMGGFGLKHPPPQIILIKNTPVPANPSTT 660
Qy 661 FSAAKFASFIQYSTGQSVSEIWELOKENSKRWNPEIQYTSNTYKNSVNVDFVTDTNGVY 720
Db 661 FSAAKFASFIQYSTGQSVSEIWELOKENSKRWNPEIQYTSNTYKNSVNVDFVTDTNGVY 720
Qy 721 SEPRPIGTRYLTRNL 735
Db 721 SEPRPIGTRYLTRNL 735

RESULT 7
Q6JC01
ID Q6JC01 PRELIMINARY; PRT; 735 AA.
AC Q6JC01;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Capsid protein VP1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388 (2004).
DR EMBL; AY530618; AAS99303.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 735 AA; 81834 MW; ACF51A69B66B582D CRC64;

Query Match 99.2%; Score 3961; DB 2; Length 735;
Best Local Similarity 98.9%; Pred. No. 1.7e-240; Mismatches 4; Indels 0; Gaps 0;
Matches 72; Conservative 4;

Qy 1 MAADGYLPDWLEDTLSEGIQWMLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFFNGLD 60
Db 1 MAADGYLPDWLEDTLSEGIQWMLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFFNGLD 60
Qy 61 KGEPVNEADAAAEHDKAYDRQLDSDGNPYLYKNHADAEOERLKEDTSFGNIGRAVFQ 120
Db 61 KGEPVNEADAAAEHDKAYDRQLDSDGNPYLYKNHADAEOERLKEDTSFGNIGRAVFQ 120
Qy 121 AKKRVLEPLGLVEPVTAPGKKRPVEHSPVEPDSSSGTGKAGQOPARKLNFGQTGDAD 180
Db 121 AKKRVLEPLGLVEPVTAPGKKRPVEHSPVEPDSSSGTGKAGQOPARKLNFGQTGDAD 180
Qy 181 SVDPDQPLGPPAPSPSGLTNTMATGSGAPMADNNEGADVGNSSGNWHCDSTWMDRVV 240
Db 181 SVDPDQPLGPPAPSPSGLTNTMATGSGAPMADNNEGADVGNSSGNWHCDSTWMDRVV 240
Qy 241 TTSTRTWALPTYNHLYKYQISSQSGASNDNHYFGYSTPWGFDYFNRFFCHFSPRDWRLI 300
Db 241 TTSTRTWALPTYNHLYKYQISSQSGASNDNHYFGYSTPWGFDYFNRFFCHFSPRDWRLI 300
Qy 301 NNNWGFPRKLNFKLFIQKVEVTONDGTITIANNLSTVQVFTDSEYQLPYVLGSAHQ 360
Db 301 NNNWGFPRKLNFKLFIQKVEVTONDGTITIANNLSTVQVFTDSEYQLPYVLGSAHQ 360
```

```
Qy 361 CLPFPFADVFMVPOQGYLTLLNGSQAVGRSSFYCLEYFPFQMLRTGNNFTFSTYFEDVPF 420
Db 361 CLPFPFADVFMVPOQGYLTLLNGSQAVGRSSFYCLEYFPFQMLRTGNNFTFSTYFEDVPF 420
Qy 421 HSSVAHSQSLDRLANPLIDQLYLYLSRTNPSTGTTTQSRLOFSQAGASDIIDQSRNWLPG 480
Db 421 HSSVAHSQSLDRLANPLIDQLYLYLSRTNPSTGTTTQSRLOFSQAGASDIIDQSRNWLPG 480
Qy 481 PCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLNVNPGPAMASHKDDKEKFFPQSGVL 540
Db 481 PCYRQQRVSKTAADNNNSEYSWTGATKYHLNGRDSLNVNPGPAMASHKDDKEKFFPQSGVL 540
Qy 541 IFGQSGSEKTNVDIEKVMITDEEIRITNPVATEQYGSVSTNLQRGNRQAATADVNTQGV 600
Db 541 IFGQSGSEKTNVDIEKVMITDEEIRITNPVATEQYGSVSTNLQRGNRQAATADVNTQGV 600
Qy 601 LPGVMWODRDVYLOGPIWAKIPHDTGHHFSPLMGGFGLKHPPPQIILIKNTPVPANPSTT 660
Db 601 LPGVMWODRDVYLOGPIWAKIPHDTGHHFSPLMGGFGLKHPPPQIILIKNTPVPANPSTT 660
Qy 661 FSAAKFASFIQYSTGQSVSEIWELOKENSKRWNPEIQYTSNTYKNSVNVDFVTDTNGVY 720
Db 661 FSAAKFASFIQYSTGQSVSEIWELOKENSKRWNPEIQYTSNTYKNSVNVDFVTDTNGVY 720
Qy 721 SEPRPIGTRYLTRNL 735
Db 721 SEPRPIGTRYLTRNL 735

RESULT 8
Q6JC21
ID Q6JC21 PRELIMINARY; PRT; 735 AA.
AC Q6JC21;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Capsid protein VP1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388 (2004).
DR EMBL; AY530598; AAS99283.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 735 AA; 81929 MW; 4EC10D74C451A429 CRC64;

Query Match 99.2%; Score 3961; DB 2; Length 735;
Best Local Similarity 99.5%; Pred. No. 1.7e-240; Mismatches 1; Indels 0; Gaps 0;
Matches 73; Conservative 1;

Qy 1 MAADGYLPDWLEDTLSEGIQWMLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFFNGLD 60
Db 1 MAADGYLPDWLEDTLSEGIQWMLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFFNGLD 60
Qy 61 KGEPVNEADAAAEHDKAYDRQLDSDGNPYLYKNHADAEOERLKEDTSFGNIGRAVFQ 120
Db 61 KGEPVNEADAAAEHDKAYDRQLDSDGNPYLYKNHADAEOERLKEDTSFGNIGRAVFQ 120
Qy 121 AKKRVLEPLGLVEPVTAPGKKRPVEHSPVEPDSSSGTGKAGQOPARKLNFGQTGDAD 180
Db 121 AKKRVLEPLGLVEPVTAPGKKRPVEHSPVEPDSSSGTGKAGQOPARKLNFGQTGDAD 180
Qy 181 SVDPDQPLGPPAPSPSGLTNTMATGSGAPMADNNEGADVGNSSGNWHCDSTWMDRVV 240
```

```
Db 181 SVDPDPLQPPAAPSGLTNTMATSGAPMADNNEGADVGNSSGNHCHDSTWMDRVI 240
Qy 241 TTSTRWALPTNNHLYKQISSQSGASNDNHFGYSTPWGYFDNRFCHFSPRDQRLI 300
Db 241 TTSTRWALPTNNHLYKQISSQSGASNDNHFGYSTPWGYFDNRFCHFSPRDQRLI 300
Qy 301 NNNWGRPKRLNFKLNFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQ 360
Db 301 NNNWGRPKRLNFKLNFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQ 360
Qy 361 CLPPFPADVFMVPOYGYLTLLNNGSOAVGRSSPYCLEYFPSSQMLRTGNNFTSYTFEDVVF 420
Db 361 CLPPFPADVFMVPOYGYLTLLNNGSOAVGRSSPYCLEYFPSSQMLRTGNNFTSYTFEDVVF 420
Qy 421 HSSYAHQSOLDRLMPLIDQVLYLRSNTTPSGTTTQSRLOFSQAGASDIRDQSRNMLPG 480
Db 421 HSSYAHQSOLDRLMPLIDQVLYLRSNTTPSGTTTQSRLOFSQAGASDIRDQSRNMLPG 480
Qy 481 PCYRQORVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDKFFPQSGVL 540
Db 481 PCYRQORVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDKFFPQSGVL 540
Qy 541 IFGKGSSEKTNVDIEKWMITDEEIRTTNPVATQYGSVSTNLQGRNQATADVNTQGV 600
Db 541 IFGKGSSEKTNVDIEKWMITDEEIRTTNPVATQYGSVSTNLQGRNQATADVNTQGV 600
Qy 601 LPGMWQDRDYLQGPWAKIPHDTGHFHPSPMGFGLKHPPIQILIKNTPVPANPSTT 660
Db 601 LPGMWQDRDYLQGPWAKIPHDTGHFHPSPMGFGLKHPPIQILIKNTPVPANPSTT 660
Qy 661 FSAAKFASFIQYSTQGVSVIEWELQKENS KRNWPEIQYTSNYSKNSVNVDFVDTNGVY 720
Db 661 FSAAKFASFIQYSTQGVSVIEWELQKENS KRNWPEIQYTSNYSKNSVNVDFVDTNGVY 720
Qy 721 SEPRPIGTRVLTNRL 735
Db 721 SEPRPIGTRVLTNRL 735

RESULT 9
Q6JC25 PRELIMINARY; PRT; 735 AA.
AC Q6JC25, 2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Capsid protein Vp1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388 (2004).
DR EMBL; AY530594; AAS99279.1; -;
DR GO; GO:0019028; C:viral capsid; IEA.
DR DR; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
DR SEQUENCE 735 AA; 81748 MW; 714DF954BF562469 CRC64;

Query Match 99.0%; Score 3955; DB 2; Length 735;
Best Local Similarity 98.9%; Pident No 4e-240;
Matches 727; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 MAADGYLPDWLEDTLSEGIQWKLKPGPPPKPAERHKDDSRGLVLPYKYLGPPFNGLD 60
|||||
Db 181 SVDPDPLQPPAAPSGLTNTMATSGAPMADNNEGADVGNSSGNHCHDSTWMDRVI 240
Qy 241 TTSTRWALPTNNHLYKQISSQSGASNDNHFGYSTPWGYFDNRFCHFSPRDQRLI 300
Db 241 TTSTRWALPTNNHLYKQISSQSGASNDNHFGYSTPWGYFDNRFCHFSPRDQRLI 300
Qy 301 NNNWGRPKRLNFKLNFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQ 360
Db 301 NNNWGRPKRLNFKLNFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQ 360
Qy 361 CLPPFPADVFMVPOYGYLTLLNNGSOAVGRSSPYCLEYFPSSQMLRTGNNFTSYTFEDVVF 420
Db 361 CLPPFPADVFMVPOYGYLTLLNNGSOAVGRSSPYCLEYFPSSQMLRTGNNFTSYTFEDVVF 420
Qy 421 HSSYAHQSOLDRLMPLIDQVLYLRSNTTPSGTTTQSRLOFSQAGASDIRDQSRNMLPG 480
Db 421 HSSYAHQSOLDRLMPLIDQVLYLRSNTTPSGTTTQSRLOFSQAGASDIRDQSRNMLPG 480
Qy 481 PCYRQORVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDKFFPQSGVL 540
Db 481 PCYRQORVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDKFFPQSGVL 540
Qy 541 IFGKGSSEKTNVDIEKWMITDEEIRTTNPVATQYGSVSTNLQGRNQATADVNTQGV 600
Db 541 IFGKGSSEKTNVDIEKWMITDEEIRTTNPVATQYGSVSTNLQGRNQATADVNTQGV 600
Qy 601 LPGMWQDRDYLQGPWAKIPHDTGHFHPSPMGFGLKHPPIQILIKNTPVPANPSTT 660
Db 601 LPGMWQDRDYLQGPWAKIPHDTGHFHPSPMGFGLKHPPIQILIKNTPVPANPSTT 660
Qy 661 FSAAKFASFIQYSTQGVSVIEWELQKENS KRNWPEIQYTSNYSKNSVNVDFVDTNGVY 720
Db 661 FSAAKFASFIQYSTQGVSVIEWELQKENS KRNWPEIQYTSNYSKNSVNVDFVDTNGVY 720
Qy 721 SEPRPIGTRVLTNRL 735
Db 721 SEPRPIGTRVLTNRL 735

RESULT 10
Q6JBZ5 PRELIMINARY; PRT; 735 AA.
AC Q6JBZ5, 2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Capsid protein Vp1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388 (2004).
DR EMBL; AY530624; AAS99309.1; -;
DR GO; GO:0019028; C:viral capsid; IEA.
GO; GO:0005198; F:structural molecule activity; IEA.
GO; GO:0019028; C:viral capsid; IEA.
```

DR	GO; GO:0005198; F:structural molecule activity; IEA.	
DR	InterPro; IPR001403; Parvo coat.	
DR	Pfam; PF00740; Parvo coat; 1.	
SQ	SEQUENCE 735 AA; 81778 MW; 15972FCAC5764274 CRC64;	
Query Match 99.0%; Score 3954; DB 2; Length 735;		
Best Local Similarity 99.0%; Pred. No. 4.6e-240;		
Matches 728; Conservative 1; Mismatches 6; Indels 0; Gaps 0;		
Qy	1 MAADGYLPDWMLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPGYKYLGPFFNGLD 60	
Db	1 MAADGYLPDWMLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPGYKYLGPFFNGLD 60	
Qy	61 KGEVNEADAAAEHDKAYDRQLDSGDNPNYLYKNHADAEFQERLKEDTSFGCNLGRAVFQ 120	
Db	61 KGEVNEADAAAEHDKAYDRQLDSGDNPNYLYKNHADAEFQERLKEDTSFGCNLGRAVFQ 120	
Qy	121 AKKRVLEPLGLVEEPVKTAQKKRPVEHSPVEPDSSSGTGKAGQOPARKRLNFQOTGDAD 180	
Db	121 AKKRVLEPLGLVEEPVKTAQKKRPVEHSPVEPDSSSGTGKAGQOPARKRLNFQOTGDAD 180	
Qy	181 SVDPDQPLGQPPAAPSGLGITNTMATGSGAPMADNNEGADGVNSGNNHCDSTWNGDRVI 240	
Db	181 SVDPDQPLGQPPAAPSGLGITNTMATGSGAPMADNNEGADGVNSGNNHCDSTWNGDRVI 240	
Qy	241 TTSTRTWTALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYDFNRRFCHFSPRDWQRLI 300	
Db	241 TTSTRTWTALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYDFNRRFCHFSPRDWQRLI 300	
Qy	301 NNNWGFPRKLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQG 360	
Db	301 NNNWGFPRKLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQG 360	
Qy	361 CLPPFPADVFWVPQGYLTLLNGSQAVGRSSFYCLEYFPSSQMLRTGNFTFSYTFEDVPP 420	
Db	361 CLPPFPADVFWVPQGYLTLLNGSQAVGRSSFYCLEYFPSSQMLRTGNFTFSYTFEDVPP 420	
Qy	421 HSSVAHSQSLDRLMNPIDQLYLRSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480	
Db	421 HSSVAHSQSLDRLMNPIDQLYLRSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480	
Qy	481 PCYRQQRVSTADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDBEKFFPQSGVL 540	
Db	481 PCYRQQRVSTADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDBEKFFPQSGVL 540	
Qy	541 IFGQSGSEKTNVDIEKVMITDEBEIRITNPVATEQYGSVSTNLQRGNRQAATADVNTQGV 600	
Db	541 IFGQSGSEKTNVDIEKVMITDEBEIRITNPVATEQYGSVSTNLQSGNTQAATSDVNTQGV 600	
Qy	601 LPGMWQDRDVLQGGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPPQILIKNTPVPANPSTT 660	
Db	601 LPGMWQDRDVLQGGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPPQILIKNTPVPANPSTT 660	
Qy	661 FSAAKFASFTQYSTGVSVSEIWELOKENSKRWNPEIQYTSNYSKSVNVDFVTDVTNGVY 720	
Db	661 FSAAKFASFTQYSTGVSVSEIWELOKENSKRWNPEIQYTSNYSKSVNVDFVTDVTNGVY 720	
Qy	721 SEPRPIGTRYLTRNL 735	
Db	721 SEPRPIGTRYLTRNL 735	
RESULT 11		
Q6JBZ9	PRELIMINARY; PRT; 735 AA.	
ID	Q6JBZ9	
AC	Q6JBZ9; 05-JUL-2004 (TrEMBLrel. 27, Created)	
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)	
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)	
DE	Capsid protein VP1.	
GN	Namescap;	
OS	Adeno-associated virus.	
OC	Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.	
OX	NCBI_TaxID=272636;	
RN	[1]	
RX	SEQUENCE FROM N.A.	
RX	PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;	
RA	Gao G., Vandenbergh L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,	
RA	Wilson J.M.;	
RT	"Clades of Adeno-associated viruses are widely disseminated in human	
RT	tissues.";	
RL	J. Virol. 78:6381-6388(2004).	
DR	EMBL; AY530620; AAS99305.1; -;	
DR	GO; GO:0019028; C:viral capsid; IEA.	
DR	GO; GO:0005198; F:structural molecule activity; IEA.	
DR	InterPro; IPR001403; Parvo coat.	
DR	Pfam; PF00740; Parvo coat; 1.	
SQ	SEQUENCE 735 AA; 81877 MW; 4CB8B93F79877D8C CRC64;	
Query Match 99.0%; Score 3953; DB 2; Length 735;		
Best Local Similarity 98.8%; Pred. No. 5.3e-240;		
Matches 726; Conservative 5; Mismatches 4; Indels 0; Gaps 0;		
Qy	1 MAADGYLPDWMLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPGYKYLGPFFNGLD 60	
Db	1 MAADGYLPDWMLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPGYKYLGPFFNGLD 60	
Qy	61 KGEVNEADAAAEHDKAYDRQLDSGDNPNYLYKNHADAEFQERLKEDTSFGCNLGRAVFQ 120	
Db	61 KGEVNEADAAAEHDKAYDRQLDSGDNPNYLYKNHADAEFQERLKEDTSFGCNLGRAVFQ 120	
Qy	121 AKKRVLEPLGLVEEPVKTAQKKRPVEHSPVEPDSSSGTGKAGQOPARKRLNFQOTGDAD 180	
Db	121 AKKRVLEPLGLVEEPVKTAQKKRPVEHSPVEPDSSSGTGKAGQOPARKRLNFQOTGDAD 180	
Qy	181 SVDPDQPLGQPPAAPSGLGITNTMATGSGAPMADNNEGADGVNSGNNHCDSTWNGDRVI 240	
Db	181 SVDPDQPLGQPPAAPSGLGITNTMATGSGAPMADNNEGADGVNSGNNHCDSTWNGDRVI 240	
Qy	241 TTSTRTWTALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYDFNRRFCHFSPRDWQRLI 300	
Db	241 TTSTRTWTALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYDFNRRFCHFSPRDWQRLI 300	
Qy	301 NNNWGFPRKLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQG 360	
Db	301 NNNWGFPRKLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQG 360	
Qy	361 CLPPFPADVFWVPQGYLTLLNGSQAVGRSSFYCLEYFPSSQMLRTGNFTFSYTFEDVPP 420	
Db	361 CLPPFPADVFWVPQGYLTLLNGSQAVGRSSFYCLEYFPSSQMLRTGNFTFSYTFEDVPP 420	
Qy	421 HSSVAHSQSLDRLMNPIDQLYLRSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480	
Db	421 HSSVAHSQSLDRLMNPIDQLYLRSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480	
Qy	481 PCYRQQRVSTADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDBEKFFPQSGVL 540	
Db	481 PCYRQQRVSTADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDBEKFFPQSGVL 540	
Qy	541 IFGQSGSEKTNVDIEKVMITDEBEIRITNPVATEQYGSVSTNLQRGNRQAATADVNTQGV 600	
Db	541 IFGQSGSEKTNVDIEKVMITDEBEIRITNPVATEQYGSVSTNLQSGNTQAATSDVNTQGV 600	
Qy	601 LPGMWQDRDVLQGGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPPQILIKNTPVPANPSTT 660	
Db	601 LPGMWQDRDVLQGGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPPQILIKNTPVPANPSTT 660	
Qy	661 FSAAKFASFTQYSTGVSVSEIWELOKENSKRWNPEIQYTSNYSKSVNVDFVTDVTNGVY 720	
Db	661 FSAAKFASFTQYSTGVSVSEIWELOKENSKRWNPEIQYTSNYSKSVNVDFVTDVTNGVY 720	
Qy	721 SEPRPIGTRYLTRNL 735	
Db	721 SEPRPIGTRYLTRNL 735	

```
RESULT 12
Q670R4
ID Q670R4 PRELIMINARY; PRT; 735 AA.
AC Q670R4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Capsid protein Vp1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen C.L., Jensen R.L., Schnepf B.C., Connell M.J., Bartlett J.S.,
RA Sferra T.J., Shell R., Johnson P.R., Clark K.R.;
RT "Characterization of Adeno-Associated Viruses in Children.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY695373; AAU05364.1; -.
DR InterPro; IPR001403; Parvo_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00740; Parvo_coat.1.
SQ SEQUENCE 735 AA; 81894 MW; ECD3D9AF395F984D CRC64;

Query Match 98.8%; Score 3946; DB 2; Length 735;
Best Local Similarity 98.5%; Pred. No. 1.5e-239;
Matches 724; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 MAADGYLPDWLEDTLSEGIQWKKLPGPPPPKPAERHKDDSGLVLPVGYKYLGPFGNGLD 60
DB 1 MAADGYLPDWLEDTLSEGIQWKKLPGPPPPKPAERHKDDSGLVLPVGYKYLGPFGNGLD 60
QY 61 KGEVPNEADAALAEHDKAYDRQLSDGDPNPKYNHADAFOERLKEBTSFGNGLGRAVQ 120
DB 61 KGEVPNEADAALAEHDKAYDRQLSDGDPNPKYNHADAFOERLKEBTSFGNGLGRAVQ 120
QY 121 AKRVLEPLGLVEBPVKTPGKKRPVEHSPVEPDSSSGTGKAGQOPARKRLNFGQTGDAD 180
DB 121 AKRVLEPLGLVEBPVKTPGKKRPVEHSPVEPDSSSGTGKAGQOPARKRLNFGQTGDAD 180
QY 181 SVDPDPLGQPPAAPSGLTNTMATGSGAPMADNNEGADGVNNGHCDSTWMDRVI 240
DB 181 SVDPDPLGQPPAAPSGLTNTMATGSGAPMADNNEGADGVNNGHCDSTWMDRVI 240
QY 241 TTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYDFNRFHCHFSPRDWQRLI 300
DB 241 TTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYDFNRFHCHFSPRDWQRLI 300
QY 301 NNNWGFPRKRLNFKLFIQVKEVTQNDGTTTANNLTSTVQVFTDSEYQLPYVLGSAHQ 360
DB 301 NNNWGFPRKRLNFKLFIQVKEVTQNDGTTTANNLTSTVQVFTDSEYQLPYVLGSAHQ 360
QY 361 CLPPFPADVFMVPOGYLTLLNGSQAVGRSSFYCLFYPFSQMLRTGNNTFSYTFEDVFP 420
DB 361 CLPPFPADVFMVPOGYLTLLNGSQAVGRSSFYCLFYPFSQMLRTGNNTFSYTFEDVFP 420
QY 421 HSSYAHQSGLDRMLNPLIDQYLYLSTNTPSGTTTQSRQLQFSQAGASDIRQSRNLWLP 480
DB 421 HSSYAHQSGLDRMLNPLIDQYLYLSTNTPSGTTTQSRQLQFSQAGASDIRQSRNLWLP 480
QY 481 PCVROQRVSKTADNNNSYSWTGATKYHLNGRDSLVPNPGPAMASHKDDDEKFPQSGVL 540
DB 481 PCVROQRVSKTADNNNSYSWTGATKYHLNGRDSLVPNPGPAMASHKDDDEKFPQSGVL 540
QY 541 IFKGQSEKTNVDIEKVMITDEEIRTNTPVATEQYGVSTNLQRNQAATADVNTQGV 600
DB 541 IFKGQSEKTNVDIEKVMITDEEIRTNTPVATEQYGVSTNLQRNQAATADVNTQGV 600
QY 601 LFGMWQDRDVLQGPWAKIPIHTDGHFHPSPIMGFGFLKHPPPQILLIKNTPVANPSTT 660
DB 601 LFGMWQDRDVLQGPWAKIPIHTDGHFHPSPIMGFGFLKHPPPQILLIKNTPVANPSTT 660
QY 661 FSAAKFASPIQTQSGVSEIWELOKENSXRNWPEIQYTSNYNKSNNVDFTVDTNGVY 720
```

```
Db 661 FSAAKFASPIQTQSGVSEIWELOKENSXRNWPEIQYTSNYNKSNNVDFTVDTNGVY 720
QY 721 SEPRPIGTRYLTRL 735
Db 721 SEPRPIGTRYLTRL 735

RESULT 13
Q6JC07
ID Q6JC07 PRELIMINARY; PRT; 735 AA.
AC Q6JC07;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Capsid protein Vp1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
RP SEQUENCE FROM N.A.
RA PubMed=15163731; DOI=10.1128/JVI.78.12.6381-6388.2004;
RA Gao G., Vandenberghe L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388 (2004).
DR EMBL; AY530612; AAS99297.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat.1.
SQ SEQUENCE 735 AA; 81661 MW; 96BB878041CB47B6 CRC64;

Query Match 98.8%; Score 3945; DB 2; Length 735;
Best Local Similarity 98.8%; Pred. No. 1.7e-239;
Matches 726; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 MAADGYLPDWLEDTLSEGIQWKKLPGPPPPKPAERHKDDSGLVLPVGYKYLGPFGNGLD 60
DB 1 MAADGYLPDWLEDTLSEGIQWKKLPGPPPPKPAERHKDDSGLVLPVGYKYLGPFGNGLD 60
QY 61 KGEVPNEADAALAEHDKAYDRQLSDGDPNPKYNHADAFOERLKEBTSFGNGLGRAVQ 120
DB 61 KGEVPNEADAALAEHDKAYDRQLSDGDPNPKYNHADAFOERLKEBTSFGNGLGRAVQ 120
QY 121 AKRVLEPLGLVEBPVKTPGKKRPVEHSPVEPDSSSGTGKAGQOPARKRLNFGQTGDAD 180
DB 121 AKRVLEPLGLVEBPVKTPGKKRPVEHSPVEPDSSSGTGKAGQOPARKRLNFGQTGDAD 180
QY 181 SVDPDPLGQPPAAPSGLTNTMATGSGAPMADNNEGADGVNNGHCDSTWMDRVI 240
DB 181 SVDPDPLGQPPAAPSGLTNTMATGSGAPMADNNEGADGVNNGHCDSTWMDRVI 240
QY 241 TTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYDFNRFHCHFSPRDWQRLI 300
DB 241 TTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYDFNRFHCHFSPRDWQRLI 300
QY 301 NNNWGFPRKRLNFKLFIQVKEVTQNDGTTTANNLTSTVQVFTDSEYQLPYVLGSAHQ 360
DB 301 NNNWGFPRKRLNFKLFIQVKEVTQNDGTTTANNLTSTVQVFTDSEYQLPYVLGSAHQ 360
QY 361 CLPPFPADVFMVPOGYLTLLNGSQAVGRSSFYCLFYPFSQMLRTGNNTFSYTFEDVFP 420
DB 361 CLPPFPADVFMVPOGYLTLLNGSQAVGRSSFYCLFYPFSQMLRTGNNTFSYTFEDVFP 420
QY 421 HSSYAHQSGLDRMLNPLIDQYLYLSTNTPSGTTTQSRQLQFSQAGASDIRQSRNLWLP 480
DB 421 HSSYAHQSGLDRMLNPLIDQYLYLSTNTPSGTTTQSRQLQFSQAGASDIRQSRNLWLP 480
QY 481 PCVROQRVSKTADNNNSYSWTGATKYHLNGRDSLVPNPGPAMASHKDDDEKFPQSGVL 540
```

```
Db 481 PCYRQVRVSKTADNNNNSEYSGTGWATKYHLNGRSLVNPMPAMASHKDDDEKFFPQSGVL 540
Qy 541 IFGQSGSEKTNVDIEKVMITDEEBIRTNVPVATEQYGSVSTNLQGRNQRAATADVNTQGV 600
Db 541 IFGQSGSEKTNVDIEKVMITDEEBIRTNVPVATEQYGSVSTNLQSGNTQAATAADVNTQGV 600
Qy 601 LPGMWQDRDVLQGPPIWAKIPHTDGHFHPSPLMGGFGLKHPPIQIILIKNTVPVANSSTT 660
Db 601 LPGMWQDRDVLQGPPIWAKIPHTDGHFHPSPLMGGFGLKHPPIQIILIKNTVPVANSSTT 660
Qy 661 FSAAKFASFTQYSTQGVSVIEWELQKENS KRNWPEIQYTSNKNKSVNVDFVDTNGVY 720
Db 661 FSAAKFASFTQYSTQGVSVIEWELQKENS KRNWPEIQYTSNKNKSVNVDFVDTNGVY 720
Qy 721 SEPRPIGTRYLTRNL 735
Db 721 SEPRPIGTRYLTRNL 735

RESULT 14
Q67006 PRELIMINARY; PRT; 735 AA.
AC Q67006;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Capsid protein VP1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
SEQUENCE FROM N.A.
RA Chen C.L., Jensen R.L., Schnepf B.C., Connell M.J., Bartlett J.S.,
RA Sferri T.J., Shell R., Johnson P.R., Clark K.R.;
RT "Characterization of Adeno-Associated Viruses In Children.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY695378; AAU05372.1; -.
DR EMBL; AY695374; AAU05366.1; -.
DR InterPro; IPR001403; Parvo coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00740; Parvo_coat_1.
SQ SEQUENCE 735 AA; 81784 MW; 818BBE497DA91984 CRC64;

Query Match 98.7%; Score 3941; DB 2; Length 735;
Best Local Similarity 98.4%; Pred. No. 3e-239;
Matches 723; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MAADGYLPDWLEDTLSGIRQWVKLKPDPKPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD 60
Db 1 MAADGYLPDWLEDTLSGIRQWVKLKPDPKPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD 60
Qy 61 KGPVNEADAALAEHDKAYDRQLDSDGNPVLKYNHADAERFQERLKEPTSGGNLGRAVFQ 120
Db 61 KGPVNEADAALAEHDKAYDRQLDSDGNPVLKYNHADAERFQERLKEPTSGGNLGRAVFQ 120
Qy 121 AKKRVLEPLGLVBEPPVKTAFGKKRPVHSVPEDSSSGTGKAGQOPARKELNFGQTGDAD 180
Db 121 AKKRVLEPLGLVBEPPVKTAFGKKRPVHSVPEDSSSGTGKAGQOPARKELNFGQTGDAD 180
Qy 181 SVDPDQPLGQPPAAPSGLTNTWATSGAPMADNEGADVGNSSGNHCDSTWMDRVI 240
Db 181 SVDPDQPLGQPPAAPSGLTNTWATSGAPMADNEGADVGNSSGNHCDSTWMDRVI 240
Qy 241 TTSTRWALPTYNHLYKQISSQSGASNDHYGYSTPWGYFDENRFCHFSRDRWQRLI 300
Db 241 TTSTRWALPTYNHLYKQISSQSGASNDHYGYSTPWGYFDENRFCHFSRDRWQRLI 300
Qy 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTIANLNTSTVQVFTDSYQLPYVLGSAHQ 360
Db 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTIANLNTSTVQVFTDSYQLPYVLGSAHQ 360
Qy 361 CLPPFPADVPVMPQYGYLTLNNGSQAVGRSSFYCLEYFPQMLRTGNFTSFYEDVPF 420
```

```
Db 361 CLPPFPADVPVMPQYGYLTLNNGSQAVGRSSFYCLEYFPQMLRTGNFTSFYEDVPF 420
Qy 421 HSSVAHQSLDRLMNPILDOVLYLSRTNTPSGTTTOSRLQFSQAGASDIRDQRNWLPG 480
Db 421 HSSVAHQSLDRLMNPILDOVLYLSRTNTPSGTTTMSRLQFSQAGASDIRDQRNWLPG 480
Qy 481 PCYRQVRVSKTADNNNNSEYSGTGWATKYHLNGRSLVNPMPAMASHKDDDEKFFPQSGVL 540
Db 481 PCYRQVRVSKTADNNNNSEYSGTGWATKYHLNGRSLVNPMPAMASHKDDDEKFFPQSGVL 540
Qy 541 IFGQSGSEKTNVDIEKVMITDEEBIRTNVPVATEQYGSVSTNLQGRNQRAATADVNTQGV 600
Db 541 IFGQSGSEKTNVDIEKVMITDEEBIRTNVPVATEQYGSVSTNLQSGNTQAATAADVNTQGV 600
Qy 601 LPGMWQDRDVLQGPPIWAKIPHTDGHFHPSPLMGGFGLKHPPIQIILIKNTVPVANSSTT 660
Db 601 LPGMWQDRDVLQGPPIWAKIPHTDGHFHPSPLMGGFGLKHPPIQIILIKNTVPVANSSTT 660
Qy 661 FSAAKFASFTQYSTQGVSVIEWELQKENS KRNWPEIQYTSNKNKSVNVDFVDTNGVY 720
Db 661 FSAAKFASFTQYSTQGVSVIEWELQKENS KRNWPEIQYTSNKNKSVNVDFVDTNGVY 720
Qy 721 SEPRPIGTRYLTRNL 735
Db 721 SEPRPIGTRYLTRNL 735

RESULT 15
Q6JC05 PRELIMINARY; PRT; 735 AA.
AC Q6JC05;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Capsid protein VP1.
GN Name=cap;
OS Adeno-associated virus.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=272636;
RN [1]
SEQUENCE FROM N.A.
RA Gao G., Vandenbergh L.H., Alvira M.R., Lu Y., Calcedo R., Zhou X.,
RA Wilson J.M.;
RT "Clades of Adeno-associated viruses are widely disseminated in human
RT tissues.";
RL J. Virol. 78:6381-6388(2004).
DR EMBL; AY530614; AAS99299.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001403; Parvo coat.
DR Pfam; PF00740; Parvo_coat_1.
SQ SEQUENCE 735 AA; 81755 MW; E1588D5235171373 CRC64;

Query Match 98.6%; Score 3940; DB 2; Length 735;
Best Local Similarity 98.9%; Pred. No. 3.5e-239;
Matches 727; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MAADGYLPDWLEDTLSGIRQWVKLKPDPKPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD 60
Db 1 MAADGYLPDWLEDTLSGIRQWVKLKPDPKPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD 60
Qy 61 KGPVNEADAALAEHDKAYDRQLDSDGNPVLKYNHADAERFQERLKEPTSGGNLGRAVFQ 120
Db 61 KGPVNEADAALAEHDKAYDRQLDSDGNPVLKYNHADAERFQERLKEPTSGGNLGRAVFQ 120
Qy 121 AKKRVLEPLGLVBEPPVKTAFGKKRPVHSVPEDSSSGTGKAGQOPARKELNFGQTGDAD 180
Db 121 AKKRVLEPLGLVBEPPVKTAFGKKRPVHSVPEDSSSGTGKAGQOPARKELNFGQTGDAD 180
Qy 181 SVDPDQPLGQPPAAPSGLTNTWATSGAPMADNEGADVGNSSGNHCDSTWMDRVI 240
Db 181 SVDPDQPLGQPPAAPSGLTNTWATSGAPMADNEGADVGNSSGNHCDSTWMDRVI 240
```

```

Db 181 SVDPQPQLGPPAAPSGLGTNTMATGSGAPWADNNEGADGVGSSGNRHCDSITWMDRVI 240
QY 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHYFGYSTPWGYFDNRFHCHFSPRDWORLI 300
Db 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHYFGYSTPWGYFDNRFHCHFSPRDWORLI 300
QY 301 NNNWGFRPKRLNFKLFNIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQG 360
Db 301 NNNWGFRPKRLNFKLFNIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQG 360
QY 361 CLPPFPADVMPVQYHYITLNGSOAVGRSSFCLEYFPPSQMLRTGNFTFSYTFEDVPF 420
Db 361 CLPPFPADVMPVQYHYITLNGSOAVGRSSFCLEYFPPSQMLRTGNFTFSYTFEDVPF 420
QY 421 HSSYAHQSGLDRNLNPLIDQVLYLSRTNTPSGTTQSRLOFSQAGASDIRDOSRNWLP 480
Db 421 HSSYAHQSGLDRNLNPLIDQVLYLSRTNTPSGTTQSRLOFSQAGASDIRDOSRNWLP 480
QY 481 PCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNPGPAMASHKXDEKEFPQSGVL 540
Db 481 PCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNPGPAMASHKXDEKEFPQSGVL 540
QY 541 IFGKQSEKTNVDIEKVMITDEEIRTNVPVATEQYGSVSTNLQRGNRQAATADVNTQGV 600
Db 541 IFGKQSEKTNVDIEKVMITDEEIRTNVPVATEQYGSVSTNLQRGNRQAATADVNTQGV 600
QY 601 LPGVMQDRDVLQGPWAKIPTHGDFHPSPLMGFGPLKHPPOILLIKNTPVPANPSTT 660
Db 601 LPGVMQDRDVLQGPWAKIPTHGDFHPSPLMGFGPLKHPPOILLIKNTPVPANPSTT 660
QY 661 FSAAKFASFTQYSTQGVSVIEWELQKENSKEWNPFIQYTSNYSKSVNVDFTVDTNGVY 720
Db 661 FSAAKFASFTQYSTQGVSVIEWELQKENSKEWNPFIQYTSNYSKSVNVDFTVDTNGVY 720
QY 721 SEPRPIGTRVLTNRL 735
Db 721 SEPRPIGTRVLTNRL 735

```

Search completed: May 19, 2005, 13:32:59
 Job time : 175 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 13:16:50 ; Search time 44 Seconds
(without alignments)
1607.257 Million cell updates/sec

Title: US-10-038-972A-13
Perfect score: 3994
Sequence: 1 MAADGYLPDWLEDTLSEGR.....TNGVYSEPRPGTGYLFRNL 735

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2548	63.8	504	1 VCPV3A	coat protein - ade
2	2122	53.1	732	2 S52210	coat protein VP1 -
3	781.5	19.6	673	1 VCPVB5	coat protein VP1 -
4	602	15.1	781	1 VCPV19	coat protein VP1 -
5	504.5	12.6	729	1 A60006	coat protein VP1 -
6	501.5	12.6	729	1 VCPVNA	coat protein VP1 -
7	461.5	11.6	727	1 VCPV1F	coat protein VP1 -
8	457.5	11.5	727	1 VCPVEP	coat protein VP1 -
9	452.5	11.3	722	1 VCPVME	coat protein VP1 -
10	442	11.1	718	1 VCPVIM	coat protein VP1 -
11	442	11.1	748	1 VCPVCP	coat protein VP1 -
12	434.5	10.9	723	1 VCPVPP	coat protein VP1 -
13	425.5	10.7	737	1 VCPVCD	coat protein VP1 -
14	411	10.3	722	1 VCPVCN	coat protein VP1 -
15	397	9.9	722	1 VCPV2	coat protein VP1 -
16	388.5	9.7	716	1 VCPV2M	coat protein VP1 -
17	232.5	5.8	587	1 B44276	coat protein VP1 -
18	213	5.3	584	2 S49594	capsid protein VP2
19	194	4.9	810	2 A44054	orf1 protein - Jun
20	170	4.3	702	1 VCPVAP	coat protein VP1 -
21	133	3.3	690	2 AB0124	probable TonB-depe
22	132.5	3.3	1129	2 T25635	hypothetical prote
23	129	3.2	743	2 T09173	EH domain protein
24	128	3.2	2529	2 B64635	toxin-like outer m
25	124.5	3.1	1145	2 T18235	transcription acti
26	122.5	3.1	1175	2 T25634	hypothetical prote
27	122	3.1	2399	2 H71879	toxin-like outer m
28	121.5	3.0	1745	2 A46431	tight junction-ass
29	121.5	3.0	2500	2 G71609	hypothetical prote

30	119.5	3.0	803	2 A86655	hypothetical prote
31	119.5	3.0	1213	2 S16356	ovo protein - frui
32	119	3.0	932	2 S62555	protoplast regene
33	119	3.0	1742	2 S76110	hypothetical prote
34	118.5	3.0	1159	2 S62562	probable nuclear p
35	118	3.0	1858	2 T18273	i-phosphatidylinos
36	117.5	2.9	2894	2 C64474	hypothetical prote
37	117	2.9	833	2 A47528	transcription fact
38	117	2.9	866	2 D44234	fibrinogen alpha c
39	117	2.9	1344	1 A35175	mucin 1 precursor,
40	115.5	2.9	309	2 I48125	syndecan-1 - Chine
41	115.5	2.9	925	2 T19361	hypothetical prote
42	115	2.9	931	2 T49710	related to glucan
43	115	2.9	1049	2 A27079	fibronectin recept
44	114.5	2.9	894	2 S51245	probable finger pr
45	114	2.9	1403	2 T11583	probable translati

ALIGNMENTS

RESULT 1

VCPV3A

coat protein - adeno-associated virus type 2

C;Species: adeno-associated virus type 2

C;Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 09-Jul-2004

C;Accession: A03698

R;Srivastava, A.; Lusby, E.W.; Berns, K.I.

J. Virol. 45, 555-564, 1983

A;Title: Nucleotide sequence and organization of the adeno-associated virus 2 genome.

A;Reference number: A03694; MUID:83164299; PMID:6300419

A;Accession: A03698

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-504 <SRI>

A;Cross-references: UNIPROT:P03135; EMBL:J01901; NID:G209616; PIDN:AAA42376.1; PID:G2096

C;Superfamily: adeno-associated virus coat protein

C;Keywords: coat protein

Query Match	63.8%	Score	2548	DB 1	Length	504			
Best local similarity	99.2%	Pred. No.	3.9e-159						
Matches	470	Conservative	1	Mismatches	1	Indels	2	Gaps	1
Qy	203	MATGSGAPMADNNEGADGVGNSGNHWCDSITWMDRVITTTSTRTWALPTYNHLYKQISS	262						
Db	1	MATGSGAPMADNNEGADGVGNSGNHWCDSITWMDRVITTTSTRTWALPTYNHLYKQISS	60						
Qy	263	QSGASNDNHYPGYSTPGWYFDNRFHCHFSPRDQORLINNNWGFPRKLNFKLFNIOVKE	322						
Db	61	QSGASNDNHYPGYSTPGWYFDNRFHCHFSPRDQORLINNNWGFPRKLNFKLFNIOVKE	120						
Qy	323	VTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPOYGYLTINN	382						
Db	121	VTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPOYGYLTINN	180						
Qy	383	GSQAVGRSSFCYLEYFPQSOMLRTGNFTFSYTFEDVPFHSSYAHSSQSLDRMLNPLIDQYL	442						
Db	181	GSQAVGRSSFCYLEYFPQSOMLRTGNFTFSYTFEDVPFHSSYAHSSQSLDRMLNPLIDQYL	240						
Qy	443	YYLRTNTPSGTTTQSRLOFSQAGASIRDOSRNWLPGPCYRQORVSKTSADNNNSYSW	502						
Db	241	YYLRTNTPSGTTTQSRLOFSQAGASIRDOSRNWLPGPCYRQORVSKTSADNNNSYSW	300						
Qy	503	TGATKYHLNGRDSLVPNGPAMASHKDEEKFPPQSGVLI1FGQKSEKTNVDIEKVMITDE	562						
Db	301	TGATKYHLNGRDSLVPNGPAMASHKDEEKFPPQSGVLI1FGQKSEKTNVDIEKVMITDE	358						
Qy	563	EEIRTTNPVATEQYGSVSTNLQRNGRQAATADVNTQGVLPQMWQDRDVIYLGQPIWAKIP	622						
Db	359	EEIGTTNPVATEQYGSVSTNLQRNGRQAATADVNTQGVLPQMWQDRDVIYLGQPIWAKIP	418						
Qy	623	HTDGHFHPPLMGFGGLKHPPPPQLIKNTVPANPSTTFSSAAKPFASFIQYSTG	676						

Db 419 HTDGHFHPPLMGFGFLKHPHPQILLIKNTVPANPSTTFSAAKFASFTIQYSTG 472

RESULT 2
S52210
coat protein VP1 - muscovy duck parvovirus
N;Alternate names: VP1 protein
C;Species: muscovy duck parvovirus
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C;Accession: S52210
R;Zadori, Z.; Erdel, J.; Nagy, J.; Kisari, J.
submitted to the EMBL Data Library, September 1993
A;Reference number: S52209
A;Accession: S52210
A;Molecule type: DNA
A;Residues: 1-732 <2AD>
A;Cross-references: UNIPROT:Q83289; EMBL:X75093; NID:G609091; PID:CAAS2984.1; PID:G609091
A;Experimental source: strain FM
C;Genetics:
A;Gene: VP1
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein

Query Match 53.1%; Score 2122; DB 2; Length 732;
Best Local Similarity 54.2%; Pred. No. 5.2e-131;
Matches 409; Conservative 92; Mismatches 194; Indels 60; Gaps 12;

QY 9 DWLEDTLSEGIQWKLKPGPPKPAE-----RHKDSRGVLVPGYKYLGFNF 57
DB 10 DWYETAAA-----SMRHLKAGAPKPKSNQOSQSVSTRKPKQRKNNRGFVLPGYKYLGFNF 65

QY 58 GLDKGEPVNEADAALAEHDKAYDRLDSDGNPNLYKYNHADAFOERLKEDTSFGNIGRA 117
DB 66 GLDKGPPVNAKADSVALEHDKAYDQQLKAGDNPYIKFKHADQEFIDNLQDTSFGNIGKA 125

QY 118 VFQAKRVLEPLGLVEHPVKTAPGK---RPVHSVP-----EPDSSSGTGKAGQOQA 167
DB 126 VFQAKRVLEPLGLVEHPVKTAPGK---RPVHSVP-----EPDSSSGTGKAGQOQA 167

QY 168 RKRLNFGQTDADSVDPDQPLQPPAPAGSLGTNTMATGSGAPMADNNEGADGVNRSNG 227
DB 186 AATEG-----SEFVAAPN-----MAGGGGAGMDAGGADGVNRSNG 223

QY 228 WHCDSTWMDRVITSTRWALPTNNHLYKQISSQSGASNDNHFYGYSTPWGYPFNR 287
DB 224 WHCDSTWMDRVITSTRWALPTNNHLYKQISSQSGASNDNHFYGYSTPWGYPFNR 287

QY 288 HCHFSPRDWORLNNHNGFPRKLNFKLKNIOUKEVTQNDGTTIANNLTSTVOVTFDSE 347
DB 284 HCHFSPRDWORLNNHNGFPRKLNFKLKNIOUKEVTQNDGTTIANNLTSTVOVTFDSE 347

QY 348 YQLPYVLSAHQGLPPFPADVFMVPOGYLTNLN---NGSQAVGRSGFYCLEYFPQMLR 404
DB 344 HOLPYVLSAHTGTPFPDSDVVALPOGYCTMHTNQSAGFNDRAFYCLEYFPQMLR 403

QY 405 TGNFTSYTFEDVPFHSVAHSQSLDRLNPLIDQVLYLSTNTTPSGTTQSRLOFSQ 464
DB 404 TGNFTSYTFEDVPFHSVAHSQSLDRLNPLIDQVLYLSTNTTPSGTTQSRLOFSQ 464

QY 465 AGASDIRDQSRNMLPGPCYQQRVSKTSADNNN-SEYS-WTGATKYHLNGRSLVNPQA 522
DB 458 AVKAGFAGMGRNMLPGKLLDQVRAYSQGTNDYANWSIWSKNKVFLLKDRYLLQPGPV 517

QY 523 MASHKODEEFPQSGVLIIFGKQ--GSEKTNVDIEKVMITDEBEIRTNVPATEQYGSVS 580
DB 518 ATTHTEQASSVPAQNIIGIAKDPYRSGSTLAGISDIWVTEQEIAPTNGVGNRPYGLTV 577

QY 581 TNLQGRNRQAATADVNTQGVLPQWMDRDLVLOGPIWAKIPHTDGHFHPPLMGFGFLK 640
DB 578 TNEQNTTATPNAELVULGALPGWQNRDIYQGGPIWAKIPKTDGKPPSNLGGFGLH 637

QY 641 HPPQILLIKNTVPANPSTTFSAAKFASFTIQYSTQSVSEIWELEKQSKRWNPFIQY 700
DB 641 HPPQILLIKNTVPANPSTTFSAAKFASFTIQYSTQSVSEIWELEKQSKRWNPFIQY 700

Db 638 NPPQVPEIKNTVPADPPLEYNQKWSYITQYSTGQCTVSMYMWELKENSKEWNPFIQF 697

QY 701 TSNYNKSNVVDFTVDYNGVYSEPRPIGTTRYLTNL 735
DB 698 TSNFGNRTSTWAPNETGGYVEDRLIGTRYLTQNL 732

RESULT 3
VCPVB5
coat protein VP1 - bovine parvovirus
N;Contains: coat protein VP2
C;Species: bovine parvovirus
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C;Accession: A26104
J;Chen, K.C.; Shull, B.C.; Moses, E.A.; Lederman, M.; Stout, E.R.; Bates, R.C.
J. Virol. 60, 1085-1097, 1986
A;Title: Complete nucleotide sequence and genome organization of bovine parvovirus.
A;Reference number: A26104; MUID:87061184; PMID:3783814
A;Accession: A26104
A;Molecule type: DNA
A;Residues: 1-673 <CHE>
A;Cross-references: UNIPROT:P07297; EMBL:M14363; NID:G333454; PIDN:AAB59847.1; PID:G80881
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein
F;138-673/Product: coat protein VP2 #status Predicted <VP2>

Query Match 19.6%; Score 781.5; DB 1; Length 673;
Best Local Similarity 29.2%; Pred. No. 2.6e-43;
Matches 213; Conservative 104; Mismatches 299; Indels 113; Gaps 22;

QY 30 PPPKPAERHKDSRGVLVPGYKYLGPFGNLDKGEPVNSADAALAEHDKAYDRLDSDGNP 89
DB 2 PPTNKA-----NSKKGLTLFPGNYLGFNSLFAGAPVNAKADAAARHKHDFGYSLLKEGKNP 57

QY 90 YLKYNHADAFOERLKEDTSFGNIGRAVFOAKRVLEPLGLVEHPVKTAPGKPEVHS 149
DB 58 YLYFNTHDQNLDELKDDTSFGKLGARVQIKKALAPAL-----POTSGKGDRLKRRK 111

QY 150 PVEPPSSSGTGKAGQOQAPARKLNFGQTDADSVDPDQPLQPPAPAGSLGTNTMATGSGA 209
DB 112 LYPARSNKAGKANREPAPSTSNQONMEVSDIPNDEAGNQ---PIELATRSV-VGSGS 166

QY 210 PMADNNEGADGVNRSNGWNCDSWMDRVITSTRWALPTNNHLYKQISSQSGASND 269
DB 167 -VGCGGRGSGVGYSTGGTGTTFSENIVTKNTRQICDKNGHLYKSEVLNTGTAH 225

QY 270 NHFYGYSTPWGYPFNRPHCHFSPRDWORLNNHNGFPRKLNFKLKNIOUKEVTQNDGT 329
DB 226 RQY-AITTPWSYFNPNQYSSHFSPNDWHLVNDYERFPRKAMIVRVYNLQIKQIMTDGAM 284

QY 330 TTIANN-LTSTVOVTFDSEYQLPYVLSAHQGLPPFPADVFMVPOGYLTNL-----NN 382
DB 285 GTVYNNDLTAGMHTPCDGDHRYPYVQHWPDDQCPNLSIWELPQYAYIPAPISVVVDN 344

QY 383 GSQAVGRS-----SPYCLEYFPQMLRTGNFTSYTFEDVPFHSVAHSQ--SLDRLMNP 436
DB 345 TINTVVEHLKGVPLMYLMSDHEVLNGRYIRYIQLWRLMDRKQHHIQHSDVOST 404

QY 437 LIDQVLYLSTNTTPSGTTQSRLOFSQAGASDIRDQSRNMLPGPCYQQRVSKTSADNN 496
DB 405 GQKQKNLLIORTKQP-----NKORFQNAAL-----RTSNMWSGF-----GIARGTHNAT 448

QY 497 NSEYSWTGATKYHLNGRSLVNPQPMASHKDDBEKFFQSGVL-----IFGKQG 546
DB 449 LQTSAGALVTMTNGAD-----VSGVRAVRVGYSTDPDIYGGQQ 487

QY 547 SEKTNVDIEKVMITDEBEIRTNVPATEQYGSVSTNLQRG--NRQAAT-----AD 594
DB 488 PE-----SDLLRLRYSAASAEGQQNPILNAAARHTFTTREATKLTGNSGADGD 536

QY 595 VNTQGVLPQWMDRDLVLOGPIWAKIPHTDGHFHPPLMGFGFLKHPPPQILLIK--NTP 652
DB 537 YKEMWMLPNQWDSAPISRSNPINWKRVRNRTLLDTQDGSIPMSHPGTFIKLARIP 596


```

QY 162 ----AGQAPARKELNFGQTGDADSDVDPDPQLPQPPAPSGGLTNTMATGSG-APMADNNE 216
DB 124 KKAGAGQV-----KKDNLAPMSDGAQVDPG---GQPAV-----RNERATGSGSGGGGGG 171
QY 217 GADGVGNSGNWH-----CDSTWMDGVRITTSRTWALTYNNHLYKQI-----S 261
DB 172 GSGGVGISTGTNNQTEPKFLENGV--EITANGSLVHLNMPSENYKRVVNNMDKTA 229
QY 262 SOSGASNDNHFGYSTPWGDFDNFRHCHFSRDPQRLINNNMFRPKRLNFKLENIOVK 321
DB 230 VKGNMALDTHVQIIVTPNPLVDANAGVWFNPGDWQLIVNTMSLHLSVFEQIFNVULK 289
QY 322 EYTON---DGTITIANNTISTVQVFTDSEYOLPVVLGSAHQGL-----PPPPA----- 367
DB 290 TVSESATQPTKVYNNDLTASLWALDSNNTPTPAAMRSETLGFYPMKPTIPTWRY 349
QY 368 ---DVFMPVQGYLTNNGSQAQVRS-----SFYCLB-YFPSSQMLRTGNRP-TFS 412
DB 350 FQWDRTLIPSH-----TGTSGTPTNIYHGTDPDDVQFYTIENSVPVHLRTGDEFATGT 403
QY 413 YTFBDVPHSSYAHSSQSLDLN--PLIDQYLYLSRTNTPS-GITTTQRLQFSQAGASD 469
DB 404 FFFDCKP--CRLTHTWQTNALGLPPLNSLPOSEGATNFGDIGVQDQKRRGVGTQMGNTD 461
QY 470 IRDQSRNLWPCPCYRQORVSKTSADNNNSSEYSGTGAHYHLNCRSLVNPGB-----A 522
DB 462 YITEATTMRP-----AEVGSAPYSPASTQ-----GPFKTPIAAG 498
QY 523 MASHKDDDEKFFPQSGVLIFGKSGSEKTNVDIE-----555
DB 499 RGAQTDENQAADGDPYAFGRGHQKTTTGETPERFYIAHODTGRYPAGDWIQNINF 558
QY 556 KMITDEEERTNPNVATQYSGVSTNLQGRNQRAATADVNTQGVLPQMWQDRDVLQ 615
DB 559 NLPTVNDNLLPTDPIG---GKTGINY-----TNFTYGLPLTALN-NVPPVTPNG 605
QY 616 PIWAKIPIHTDGHFHPSPLMGGFGLKHP-----PQILLIKNTVPANPSTTSAKFAS 668
DB 606 QIWDKEFTD-----LKPRLHVNAPFVCQNCQGLFVKVAPNLNNEYDPPDASAMSR 658
QY 669 FITQYSTQGVSEIWELEQKENSKNWPEIQYTSNKNKSNVND 711
DB 659 IVT-YSDFWWKGKLVFKAKLRASHTNP-IQ-----QMSINVD 694

RESULT 10
VCPVIM
coat protein VP1 - minute virus of mice (strain MMV1)
C:Species: minute virus of mice, murine parvovirus
C>Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 09-Jul-2004
C:Accession: B23008; B29510
R:Sabli, R.; McMaster, G.K.; Hirt, B.
Nucleic Acids Res. 13, 3617-3633, 1985
A:Title: DNA sequence comparison between two tissue-specific variants of the autonomous
A:Reference number: A23008; MUID:85242059; PMID:3855242
A:Accession: B23008
A:Molecule type: DNA
A:Residues: 1-718 <SAH>
A:Cross-references: UNIPROT:P07302; EMBL:X02481; MID:G60918; PIDN:CAB46507.1; PID:G54198
R:Ascoli, C.K.; Gardiner, E.M.; Tattersall, P.
J. Virol. 57, 656-669, 1986
A:Title: DNA sequence of the lymphotropic variant of minute virus of mice, MMV(1), and
A:Reference number: A29510; MUID:86115415; PMID:3502703
A:Accession: B29510
A:Molecule type: DNA
A:Residues: 1-143, A', 145-718 <AST>
A:Cross-references: EMBL:M12032
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein

Query Match 11.1%; Score 442; DB 1; Length 718;
Best Local Similarity 24.8%; Pred. No. 4.7e-21;

```

```

Matches 201; Conservative 99; Mismatches 294; Indels 218; Gaps 38;
QY 45 LVLPGYKYLGPNGLDGEPVNEADAAALBHDKAYDRQLDSDGNPYLYKNHADAEPQERL 104
DB 1 MVPFGYKYLGPNGLDGEPVNEADAAALBHDKAYDRQLDSDGNPYLYKNHADAEPQERL 104
QY 105 KEDTSPGNTGRVAFQAKRVLPLGLVPEPVKTA PGKKRPVEHSPVPSDSSGSGTKAG- 163
DB 61 KDARDMGKGVGHVFFRTKRAFA PKLATDSEP-----GTSVGRAGK 101
QY 164 -----QOPARKEL-----NFGQT-GDADSVDPDPQLPQPPAPSGGLTNTMAT 205
DB 102 RTRPPAYIFINQARAKKLTSAAQQSSQTSMDGTSQPD-----CGGVHSAAR 150
QY 206 -----GSGMAPDNNEGADGVNSSGNHCHDS--TWMGD-----RVITTSRTWAL-- 249
DB 151 VERAADGPGSGG-----GSGGGGVGVSTGSDNTHYRFLGDDGWVEITALATLVLNM 206
QY 250 PTYNN-----HLYKQISSQSGASNDNHFGYSTPWGDFDNFRHCHFSRDPQRLINNN 303
DB 207 PKSENYCRIRVNTTDTSVKGNMAKDDAHEQIWTWPSLVDANAGVWLQPSDMQYICNTM 266
QY 304 WGRPRKRLNFKLNIQVKEVTQND-GTTTIA-----NNLTSTVQVFTDSEYOLPVVLGSAHQ 359
DB 267 SOLNLVSLDQEIFNVVLKTVTTEQDSGQAIKYNDLITACMMVAVDSSNNILPYTPAANS 326
QY 360 GCLPPFPADVFMVPOGY-----LTLNN-----GSQAVGRSSFYCLEYFPS 400
DB 327 ETLGFPYKPKTIASPYRYFCVDRDLSVTYENQGTIEHNVMGTPKGMNSQFTIENTQQ 386
QY 401 -OMLRGNRP--TFSYTFEDVPFHSYAHSSQSLDLN--PLIDQYLYLSRTNTPSGTGT 456
DB 387 ITLLRTGDEFATGTYYFTNPV--KLTHWTQNRQLGQPLLSTF-----PEADTDAGTLT 440
QY 457 QSRLOFSQAGASDIRDOSRWL-----PGCYRQQRVS-----KTSAD-- 494
DB 441 ---AQGRHGHATQ---EVNVVSEAIRTPAQVGFQCPHNDFEASRAGFAAPKPADVT 494
QY 495 -----NNNSEYSWTGATKYHLNCRDLSLVNPPAPAMASHKDDDEK------PQSG 538
DB 495 QGVDRANGSVYSY---GKQH---GENWAAGPAPERWTWDETNPFGSGRDRDGFIOQA 548
QY 539 VLIFGKQSGSEKTNVLEKWMITDEEIRITNPVATEQYSGVSTNLQGRNQRAATADVNTQ 598
DB 549 PLVPPPLNG-----ILTNPANPDKNDI---HFSNV-----FNSY 581
QY 599 GVLPGMWQDRDYLQGPWAKIPHTDGHFHPSPLMGGFGLKHP-----PQILLIKNT 652
DB 582 GLPTAFS-HPSPYVQGIWDK-----ELDLHKRPLRHITAPFVCKKNAP 625
QY 653 ---VPANPSTTF---SAAKFPASITQYSTQGVSEIWELEQKENSKNWNP--EIQVTS 702
DB 626 GQMLVRLGLNLTQDQDPNGATLSRIVT-YGTFPFWKGLTMRKLRANNTWNPVQVSVED 684
QY 703 NYNKSNNVDFTVDTNGVSEPRPIGTRYLTRN 734
DB 685 NGNSYMSVTWKLPTATGNQSVLITRPAEN 716

RESULT 11
VCPVCP
coat protein VP1 - canine parvovirus (strain N)
N:Contains: coat protein VP2
C:Species: canine parvovirus, CPV
C>Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
C:Accession: B29962
R:Reed, A.P.; Jones, E.V.; Miller, T.J.
J. Virol. 62, 266-276, 1988
A:Title: Nucleotide sequence and genome organization of canine parvovirus.
A:Reference number: A29962; MUID:88062992; PMID:2824850
A:Accession: B29962
A:Molecule type: DNA
A:Residues: 1-748 <REE>

```


A;Cross-references: UNIPROT:P12930; EMBL:M19296

C;Accession: B33302

A;Introns: 26/3

C;Superfamily: parvovirus coat protein

C;Keywords: coat protein

F;155-748/Product: coat protein VP2 #status predicted <VP2>

Query Match 11.1%; Score 442; DB 1; Length 748;
Best Local Similarity 24.2%; Pred. No. 5e-21;
Matches 194; Conservative 99; Mismatches 306; Indels 202; Gaps 35;

```

Qy 30 PPPKPAERHK-----DDSR-----GLVLPGYKYLGFNGLDKGEFVNEADAAALE 74
Db 3 PPPKPAERHGKVLVKWEGKDLITMCFFIGLVPFGYKYLGFNGSLDGEFTNPDSAAAKE 62
Qy 75 HDKAYDRQLDGDNPYLKYNHADAEFOERLKEDTSFGGNLGRAVFOAKKRVLEPL-GLVE 133
Db 63 HDEAYAYLRSGKNPYLYFSPADQRFIDQTKADKGGKIGHYFFRAKATAPVLDTDPD 122
Qy 134 EPVKTAFGKRPVBHSPVEPDSSGTGK-----AQQPARKRLNFGQTGDADSDVPDPQL 188
Db 123 HPSTSRPTK--PTKRSPPPHIFINLAKKKKAGAGQV---KRDNLAPMSDGAQVDPG--- 174
Qy 189 GPPAAPSGLTMTATGSG-APMADNNEGADGVNSSGNWH-----CDSTWMDRVI 240
Db 175 GQPAV-----RNERATGSGSGGGGGGGVGISTGTNNQTEFKFLENGWV--EIT 226
Qy 241 TTSRTTALPTYNHLYKQI-----SSQSGASNDNHFGYSTPWPYDFNFRFCHFS 292
Db 227 ANSRLVHLNMPSENRRVVVNNMDKTAVNGNMLDIDHAIQVTPSLVLDANAGWVFN 286
Qy 293 PRDQRLINNMGFRPKRLNFKLNIQVKEVTQN---DGTITANNLTSTVQVFTDSEYQ 349
Db 287 PGDWQLIVNTMSLHLSVFEQIEINVLKTVSEATQPTKYVNDLTASLWALDSNNT 346
Qy 350 LPYVLGSAHQCL-----PPFPA-----DVFMVPOYGYLTINNGSQVGRS----- 390
Db 347 MPFTPAAMRSETLGFYPMKPTIPTWRYFYFQWDRTLIPSH-----TGTSPTPTNIYHGT 400
Qy 391 -----SFYCLE-YPPSOMLRTGNFTSYTPEDV-----LMNPLIDQYL--YYLSRTNTPSGTTQ 457
Db 401 DPDDVQFYTIENSVPVHLLRTGDSEFATGTFDFDCKPCRLTHWTQNRALGPLPENSUPQ 460
Qy 427 SQSLDR-----LMNPLIDQYL--YYLSRTNTPSGTTQ 457
Db 461 SEGATNFGDIGVQDQKRGVQTMGNTNVTIETIMRPAEVCYSAPYYSFEASTQGPFTK 520
Qy 458 SRLQFSQAGASDIRDQSRNWLPGCYRQQRVSKTSADNNNSEYSWTGATKHLNGRDSLV 517
Db 521 --IAAGRGGAQTYENQAADGDPYAFGRQHGXQKTTT-----TGSTPERF----- 562
Qy 518 NPGPAMASHKDDEKFPQSGVLIFGKGSEKTNVDIEKVMITDDEEIRTNVATBQYG 577
Db 563 -----TYIAHQDTGR--YF-----EGOWIQINIF-NLPVTNDNVLLPTDPIG-----G 602
Qy 578 SVSTNLQGRNQAATADVNTQGLPFGVMWDRDYLQGPVWAKIPHTDGHFHPSPLMGGF 637
Db 603 KTGINY-----TNIFNTYGLTALN-NVPPVYPNGQIWDKEFTD-----LKPL 646
Qy 638 GLKHP-----PPQIILKNTVPANSTPTSAKFAFSPITQYTGQVSVEIEWELOKEN 690
Db 647 HVNAPFVQNNCPQLFVKAPNLNTEYDPPDASANMSRIVT-YSDFWWKGKLVFAKLRA 705
Qy 691 SKRWNPETQYTSNYSNVND 711
Db 706 SHTWNP-IQ-----QMSINVD 720

```

RESULT 12

VCPVPP

coat protein VP1 - porcine parvovirus

N;Contains: coat protein VP2

C;Species: porcine parvovirus

C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C;Accession: B33302
R;Ranz, A.I.; Mancius, J.J.; Diaz-Aroca, E.; Casal, J.I.
J. Gen. Virol. 70, 2541-2553, 1989
A;Title: Porcine parvovirus: DNA sequence and genome organization.
A;Reference number: A33302; MUID:90010964; PMID:2794971
A;Accession: B33302
A;Molecule type: DNA
A;Residues: 1-723 <RAN>
A;Cross-references: UNIPROT:P18546; EMBL:D00623
C;Genetics:
A;Introns: 10/1
C;Superfamily: parvovirus coat protein
C;Keywords: coat protein
F;145-723/Product: coat protein VP2 #status predicted <VP2>

Query Match 10.9%; Score 434.5; DB 1; Length 723;
Best Local Similarity 25.1%; Pred. No. 1.5e-20;
Matches 199; Conservative 107; Mismatches 318; Indels 169; Gaps 37;

```

Qy 30 PPPKPAERHKDDSRGLVLPYKYLGFNGLDKGEFVNEADAAALEHDKAYDRQLDSDGNP 89
Db 3 PPPKPA-----RGTI-----LPGNSLDQGEFTNPSDAAAKEHDEAYDKYIKSGNP 49
Qy 90 YLKNHADAEFOERLKEDTSFGGNLGRAVFOAKKRVLEPLGLVBEVPKTA PGKRPVHVS 149
Db 50 TFVPSAADEKFIKETEHAKDYGGKIGHYFFRAKAFAPKLSETDSPTTS---QQPEVRS 106
Qy 150 PVEPDSSSGTGK-----AQQPARKRLNFGQTGDADSDVPDPQLGPPAPPSGLG 199
Db 107 PRKHGSKPPKRPAPRHIFINLAKKKAKGTSNTNSMSNVEQHNPI-----NAA 158
Qy 200 TMTATGCS---GAPMADNNEGADGVNSSGNWH---CDSTWMD---RVITSTRTALPT 251
Db 159 TEISATNGESGGGGGGGAGGVGVTSGFNQTEFVYLGEGVLVITAHASLIHLAM 218
Qy 252 YNNHLYKQI---SSQSGAS---NDNHFGYSTPWPYDFNFRFCHFSPPRDWQRLNNW 304
Db 219 PEHETVKRIHVLNSESAGQWQDDAHTQVTFWLSLIDANAGWVFNPAWQLISNMT 278
Qy 305 GFRPKRLNFKLNIQVKEVTQN---DGTITANN--LTSTVQVFTDSEVQLPYVLSAHQGC 361
Db 279 EINLVSFEQAFINVLKTIETESATSPPTKIYNDLTASLWALDNTNLTPTPAAPRSET 338
Qy 362 LPPPPADVFMVPOYGYL-----TLNNGSQVGRSS-----FYCLE-YPPSQ 401
Db 339 LGFYFWLPTQTYRYVYLSICIRNLNPPYTYGSGQPNRLNRLHSDIMFTIENAVPIH 398
Qy 402 MLRTGNF--TFSYTFEDVPFHSSYAHQSQSLDR-----LMNPLI--DQYLYLSRTNT 450
Db 399 LLRTGDEFSTGIYHFDTKPL--KLTHSQWTRSLGLPPKLLTEPTTEGDQHPGTLPAANT 456
Qy 451 PSGTTTQSRLOFSQAGASDIRDQSRNWLPGCYRQQRVSKTSADNNNSEYSWTGATKYL 510
Db 457 RKGHYHOTINNSYTEATA--IRPAQVGY-----NTPYMNFEYS----- 491
Qy 511 NGRDNLNPGPAMASHKDDEKFPQSGVLIFGKGSEKTNVDIEKVMIT---DDEELRT 567
Db 492 NGGPFLTPVPTADTQYNDDE-----PNGAIRFTMDYQHGHLTSSQELERYT 539
Qy 568 TNP-----VATEQYG-SVSTNLQGRNQAATAD-----VNTQGVLPGMVW 606
Db 540 FNPQSKGRAPKQKQFNQQAFLNLENTNNGTLTLPDPIGGKSNHFMFTLNTYGLTALN- 598
Qy 607 QDRDVLQGPVWAKIPHTD--GHFHPSPLMGGFGLK--HPPQIILKNTVPANSTPTFSA 663
Db 599 NTAPVFPNGQIWDKELDTDLKPLH---VTAPFYCKNPNPQQLFVKIAP---NLTDFFNA 652
Qy 664 -AKFASPIQYTSQVSVEIEWELOKENSKRWNPETQYTSNYSNVNDVTDVNTN--GV- 719
Db 653 DSPQOPRIITYSNFWKGTLTFTAKMRSSNNWNPVIOQHTTT---AENIGKIPTNIGIK 709
Qy 720 ----YSE--PRPI 726

```

Db 710 MFPEYSOLIPRKL 722

RESULT 13

VCPCVD

N/Contains: coat protein VP2

C/Species: canine parvovirus, CPV

C/Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004

C/Accession: A31163

R/Parish, C.R.; Aquadro, C.F.; Carmichael, L.E.

Virolgy 166, 293-307, 1988

A/Title: Canine host range and a specific epitope map along with variant sequences in the

A/Reference number: A31163; MUID:89020796; PMID:3176341

A/Accession: A31163

A/Molecule type: DNA

A/Residues: 1-737 <P>

A/Cross-references: UNIPROT:P17455; EMBL:M23255; NID:g333467; PIDN:AAA47158.1; PID:g3334

C/Genetics:

A/Introns: 26/3

C/Superfamily: parvovirus coat protein

C/Keywords: coat protein

F:584-737/Product: coat protein VP2 #status predicted <VP2>

Query Match 10.7%; Score 425.5; DB 1; Length 737;

Best Local Similarity 24.4%; Pred. No. 5.9e-20;

Matches 190; Conservative 102; Mismatches 317; Indels 171; Gaps 34;

QY 30 PPPPAERHKDD----SRGLVLPGVKYLGPNGLDKGEPVNEADAAALEHDKAYDRQLDS 85

DB 3 PPAKRRARGKVLVKWGEKDLITYKVLGPNSLDQGEPTNPSDAAAKEHDEAVAAAYLRS 62

QY 86 GDNPLKYNHADAFQERLKEDTSGNGLGRAVFAQKRVLEPL-GLVEEPVKTA PGKKR 144

DB 63 GKPYLTFSPADQFIDQTKDAKDGWKIGHYFFRAKKAIAPLVLTDTDPHSTSRPTK-- 120

QY 145 PVEHSPVEPPSSSGTGK----AQQPARKRLNGQTGDADSDVPDQPLGPPPAAPSGLG 199

DB 121 PTKRSKPPPHIFINLAKKKKAGQV---KRDNLAPMSDGAVQPDG---GQPAV----- 168

QY 200 TTNMTATGSG-APMADNNEGADGVNSSGNWH-----CDSTWMDRVITSTRTWALPT 251

DB 169 RNERATGSGNGSGGGGGGGVIGTGTENNQTEFKFLENGWV--BITANSSRLVHLNM 226

QY 252 YNNHLYKQI-----SSQSGASNDNHFGYSTPMGYEDFNRFCHFSPRDWORLNNN 303

DB 227 PSENYRVRVVVNMMDKATVNGNMLDDIHAQIVTFPSLVDANAGWVFNPDGMDQLIVNTM 286

QY 304 WGFPRKRLNFKLNIOVKVTQN---DGTITIANNTTSTVQVFTDSEYQLPYVLGSAHQ 360

DB 287 SELHLVSFQEIFNVVLKTVSESATOPPTKVYNNDLTASLWALDNNNTMPTFAAMRSE 346

QY 361 CL-----PPPPA-----DVFMVPOQVGYLTNNGSQAQVRS-----SPCYLE 396

DB 347 TLGFYFPWKPTIPTPRYYPQWDRTLPSH-----TGTSGPTNIYHGTDPODDVQYTTIE 400

QY 397 -YFPQMLRTGNF--TFSTFEDVFPHSHYAHQSGLRLMNLPLIDQYLXLSRTNTPSGT 454

DB 401 NSVPVHLRLTGDFEATGFFFDCKP--CRLHTWQTNALG--LPPFLNLSLQPS---EGA 453

QY 455 TTQSRLOPSQAGASDIRDQSRNLWPGCYRQQRVSKTADNNNSE--YSWTGATKVHLNG 512

DB 454 T-----NFGDIGVQODKRGVTVQMGNTNVTIETATMRPAEVCYSAPYVSFEASTQ----- 503

QY 513 RDSLNVNPGP-----AMASHKDDEEKFFPQSGVLI FGKQSEKTNVDIE----- 555

DB 504 -----GPFKPTIAGGGGAQTDBENQAADGNPRYAFGRQHQKTTTGTTPERTYIAH 556

QY 556 -----KVMITDEEIRTTNPVATEQYGSVSTNLQRNQRQAATADVNTQ 598

DB 557 QDTRYPEGDWIQINFNLPVTDNVLPTDPIG-----GKTGINY-----TNIFTY 604

```

556 -----KVMITDEEBEIRTNTPVATEQYGSVSTNLQRGRQAATADVNTQGVLPGMV 605
      : : : : :
549 EGDWIONINFNLPTVNDNVLLPIDIG-----SKTGINY-----TNINFTYGPLTALN 596
      : : : : :
606 WQDRVYVYLGPTWAKIPHTDGHFHFSPLMGGFGLKHP-----PPQILIKNTPVPANPS 658
      : : : : :
597 -NVPPVYVNGQIWDKEFDTD-----LKPRLHVNAFPVCQNNCPQLFVKLAPNLNEY 648
      : : : : :
659 TTFSAAKASFITQYSTGCVSVSEIWELEQKENSKEWNPBEIQYTSYNKSVNVD 711
      : : : : :
649 DFDASANNRSRIVT-YSDFWWKGKGVFKAKLRASHTWNP-IQ-----QMSINVD 694
      : : : : :

```

RESULT 15

VCPVV2
coat protein VP1 - parvovirus H1
C:Species: parvovirus H1
A:Note: host Homo sapiens (man)
C:Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 09-Jul-2004
C:Accession: A03699
R:Rhode III, S. L.; Paradiso, P. R.
J. Virol. 45, 173-184, 1983
A:Title: Parvovirus genome: nucleotide sequence of H-1 and mapping of its genes by hybridization
A:Reference number: A03695; MUID:83112183; PMID:6823009
A:Accession: A03699
A:Molecule type: DNA
A:Residues: 1-722 <RHO>
A:Cross-references: UNIPROT:P03136; EMBL:X01457; EMBL:J02198
C:Superfamily: parvovirus coat protein
C:Keywords: coat protein

Query Match	9.9%	Score 397;	DB 1;	Length 722;
Best Local Similarity	23.7%	Fred. No. 4.2e-18;		
Matches 187;	Conservative 110;	Mismatches 325;	Indels 168;	Gaps 37;
Qy	30	PPPKPAERHKDDSRGLVLP	PGYKYLPGFNGLDKGPVNEADAAA	LEHDKAYDRQLDSDGDP 89
Db	3	PPAKRARG-----	-----NSLDQGEFTWPSDAAAKEHDEAYDQYIKSGNKP 44	
Qy	90	YLYKNHADAEOERLEKEDTSPG	NLGRAVFOAKKRVLEPLGLVBEVPKTA	---PGKK-RP 145
Db	45	YLYESPADQRIDQTKADKMGW	KGVGHYFFRTKRAFAPKLSTDSEPGT	SGVSRPGKRTKP 104
Qy	146	VEHSPVEPSSSGTKACQOQ	PARXLNFGQTGDADSVDPDPOPLQCP	PAAPSGELGNTWAT 205
Db	105	PAHIFVHQ-----	---ARAKKRSALAAQORTLMTSDGTET	NPQDTGIANARVERSAD 154
Qy	206	GS GAPMADNNEGADGVGNSSG	NHWCDSW---MGRVITTTSTRT-----	WALPTVNNHLYK 258
Db	155	GGGS-SGGSGSGGGIGVSTG	TYDNQTYIKELGDGWEITAHASRLHL	HGNPPSENCRV 213
Qy	259	QISSQSGASNDNHVFG---YS	-----TPWCYGFNRPCHFSPRDMOR	LINNNGWRFRPK 309
Db	214	TVHNNQTTGHGTVKVGKNMAY	THQOIWTPWSLVDANAGVWFOPSDW	OFIQNSHESLNLD 273
Qy	310	RLNFKLNIQVKEVTONDGTIT	IA-----NNLTSTVQVFTDSEVQL	PVVLGSAHOGCL-- 362
Db*	274	SLSQELFNVVVKTITTEQGAG	QDAIKVYNNDLTACMMVALDSNNIL	FYPAAQTSSETLGF 333
Qy	363	----PPFPAD---VFMPQY	LYLTNLNGSQ-----AVG-----	RSSPFCYLE-YFPSONML 403
Db.	334	YPWKPTAPAPVRYFFWPRQL	SVTSNNSAEGTQITDIGEPOALNSQ	FTIENTLPTILL 393
Qy	404	RTGNNTFP-SYTFPDEVPHSS	YA-----HSOSLRLMN-PLIDQV	LYVLSRNTNTPSGTT- 455
Db	394	RTGDEFTGTVIFNTDPLKLT	HTHTWQTNRHLCACLOGITDLP	STSDATASLTANGDRFSGTQ 453
Qy	456	TQS-----RLQFSQAGAS	DIIDSRNWLPGCYRQORVSKTSAD	-NNNSEYSWTGAT 506
Db	454	TQNVNYVTEALRTPAQIG	FQMOPHNFANRGGFPF-----	KVPVPLDITAGEDHDANGAI 509
Qy	507	KYHLNGRDSLNVPGPAMASH	KDDDEKFFPQSGVLIFGQKQSEK	TNVDIEKWIITDEBEIR 566

[illegible]

Search completed: May 19, 2005, 13:33:49
Job time : 47 secs

A: Title: Parvovirus genome: nucleotide sequence of H-1 and mapping of its genes by hybridization
A: Reference number: A03695; NUID:83112183; PMID:6823009
A: Accession: A03699
A: Molecule type: DNA
A: Residues: 1-722 <RHO>
A: Cross-references: UNIPROT:P03136; EMBL:X01457; EMBL:J02198
C: Superfamily: parvovirus coat protein
C: Keywords: Coat protein

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 13:33:06 ; Search time 139 Seconds
(without alignments)
1768.799 Million cell updates/sec

Title: US-10-038-972A-13

Perfect score: 3994

Sequence: 1 MAADGYLPDWLEDTLSEGR.....TNGVSEPRPTGTRYLTRLN 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1434725 seqs, 334507595 residues

Total number of hits satisfying chosen parameters: 1434725

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:**

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pcp.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pcp.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pcp.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pcp.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3994	100.0	735	13	US-10-038-972A-13
2	3994	100.0	735	14	US-10-293-478-1
3	3994	100.0	735	14	US-10-291-583-70
4	3994	100.0	735	15	US-10-423-704A-4
5	3994	100.0	735	17	US-10-496-799-4
6	3989	99.9	735	17	US-10-880-297-11
7	3985	99.8	735	17	US-10-880-297-15
8	3984	99.7	735	17	US-10-880-297-5
9	3984	99.7	735	17	US-10-880-297-13
10	3984	99.7	735	17	US-10-880-297-19
11	3983	99.7	735	17	US-10-880-297-7
12	3980	99.6	735	17	US-10-880-297-9
13	3980	99.6	735	17	US-10-880-297-31

14	3975	99.5	735	17	US-10-880-297-25	Sequence 25, Appl
15	3975	99.5	735	17	US-10-880-297-27	Sequence 27, Appl
16	3969	99.4	735	17	US-10-880-297-17	Sequence 17, Appl
17	3969	99.4	735	17	US-10-880-297-39	Sequence 39, Appl
18	3962	99.2	735	17	US-10-880-297-23	Sequence 23, Appl
19	3962	99.2	735	17	US-10-880-297-41	Sequence 41, Appl
20	3944	98.7	735	17	US-10-880-297-35	Sequence 35, Appl
21	3941	98.7	735	17	US-10-880-297-21	Sequence 21, Appl
22	3936	98.5	735	17	US-10-880-297-37	Sequence 37, Appl
23	3930	98.4	735	17	US-10-880-297-29	Sequence 29, Appl
24	3929	98.4	735	17	US-10-880-297-33	Sequence 33, Appl
25	3623	90.7	735	14	US-10-291-583-67	Sequence 67, Appl
26	3622	90.7	735	14	US-10-291-583-69	Sequence 69, Appl
27	3617	90.6	735	14	US-10-291-583-68	Sequence 68, Appl
28	3615	90.5	735	14	US-10-291-583-66	Sequence 66, Appl
29	3532.5	88.4	736	14	US-10-291-583-71	Sequence 71, Appl
30	3532.5	88.4	736	15	US-10-423-704A-6	Sequence 6, Appl
31	3532.5	88.4	736	17	US-10-496-799-6	Sequence 6, Appl
32	3421	85.7	731	14	US-10-291-583-88	Sequence 88, Appl
33	3416.5	85.5	738	14	US-10-291-583-91	Sequence 91, Appl
34	3409.5	85.4	736	14	US-10-291-583-65	Sequence 65, Appl
35	3405.5	85.3	738	14	US-10-291-583-93	Sequence 93, Appl
36	3405.5	85.3	738	14	US-10-291-583-94	Sequence 94, Appl
37	3404.5	85.2	738	14	US-10-291-583-79	Sequence 79, Appl
38	3404.5	85.2	738	14	US-10-291-583-81	Sequence 81, Appl
39	3404.5	85.2	738	14	US-10-291-583-85	Sequence 85, Appl
40	3402.5	85.2	736	14	US-10-291-583-64	Sequence 64, Appl
41	3402.5	85.2	736	15	US-10-423-704A-5	Sequence 5, Appl
42	3402.5	85.2	736	15	US-10-696-261-3	Sequence 3, Appl
43	3402.5	85.2	736	15	US-10-696-261-13	Sequence 13, Appl
44	3402.5	85.2	736	15	US-10-696-282-3	Sequence 3, Appl
45	3402.5	85.2	736	15	US-10-696-282-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-10-038-972A-13
; Sequence 13, Application US/10038972A
; Publication No. US20020192823A1
; GENERAL INFORMATION:
; APPLICANT: J. Bartlett
; TITLE OF INVENTION: AAV VECTORS AND METHODS
; FILE REFERENCE: 28335/36996US
; CURRENT APPLICATION NUMBER: US/10/038,972A
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: US 60/260,124
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 735
; TYPE: PRT
; ORGANISM: adeno-associated virus 2 VLP capsid protien
US-10-038-972A-13

Query Match 100.0%; Score 3994; DB 13; Length 735;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAADGYLPDWLEDTLSEGIQWKLKPGPPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD	60
Db	1	MAADGYLPDWLEDTLSEGIQWKLKPGPPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD	60
Qy	61	KGPEVNEADAALAEHDKAYDRQLDSDGNPNYLYKNHADAERLKDFTSFCGNLGRAVFQ	120
Db	61	KGPEVNEADAALAEHDKAYDRQLDSDGNPNYLYKNHADAERLKDFTSFCGNLGRAVFQ	120
Qy	121	AKKRVLEPLGLVVEPVKTAPCKKPVHSVPEPPSSSTGKAGQOPARKRLNFGQTGDAD	180
Db	121	AKKRVLEPLGLVVEPVKTAPCKKPVHSVPEPPSSSTGKAGQOPARKRLNFGQTGDAD	180

181 SVDPDQPLGPPAAPSGLTNTMATGSGAPMADNNEGADGVGNSGNWHCDSTWMDRVI 240
Db |||||
181 SVDPDQPLGPPAAPSGLTNTMATGSGAPMADNNEGADGVGNSGNWHCDSTWMDRVI 240
Qy |||||
Db |||||
241 TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPWGYPDFNRFCHFSPRDWORLI 300
Db |||||
241 TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPWGYPDFNRFCHFSPRDWORLI 300
Qy |||||
Db |||||
301 NNNWFRPKRLNFKLNFQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQG 360
Db |||||
301 NNNWFRPKRLNFKLNFQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQG 360
Qy |||||
Db |||||
361 CLPPFPADVFWVPOQYGYLTLLNGSQAAGRSSFYCLEYFPQSOMLRTGNNFTSYTFEDVPF 420
Db |||||
361 CLPPFPADVFWVPOQYGYLTLLNGSQAAGRSSFYCLEYFPQSOMLRTGNNFTSYTFEDVPF 420
Qy |||||
Db |||||
421 HSSYAHQSQSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOFSQAQASDIRDQSRNMLPG 480
Db |||||
421 HSSYAHQSQSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOFSQAQASDIRDQSRNMLPG 480
Qy |||||
Db |||||
481 PCYRQORVSKTSADNNSEYSGWKATKYHLNGRDSLVPNGPAMASHKDDKEKFFPQSGVL 540
Db |||||
481 PCYRQORVSKTSADNNSEYSGWKATKYHLNGRDSLVPNGPAMASHKDDKEKFFPQSGVL 540
Qy |||||
Db |||||
541 IFKGQSEKTNVDIEKVMITDEEIRITNPVATEQYGSVSTNLQGRNQAATAADVNTQGV 600
Db |||||
541 IFKGQSEKTNVDIEKVMITDEEIRITNPVATEQYGSVSTNLQGRNQAATAADVNTQGV 600
Qy |||||
Db |||||
601 LPMWQORDVYLOQPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQIILIKNTVPANPSTT 660
Db |||||
601 LPMWQORDVYLOQPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQIILIKNTVPANPSTT 660
Qy |||||
Db |||||
661 FSAAKFASFIQYSTQGVSVIEWELOKENSKRNPEIQYTSNKNKSNVNDFTVDTNGVY 720
Db |||||
661 FSAAKFASFIQYSTQGVSVIEWELOKENSKRNPEIQYTSNKNKSNVNDFTVDTNGVY 720
Qy |||||
Db |||||
721 SEPRPIGTRYLTRNL 735
Db |||||
721 SEPRPIGTRYLTRNL 735

RESULT 2

US-10-293-478-1
; Sequence 1, Application US/10293478
; Publication No. US20030078411A1
; GENERAL INFORMATION:
; APPLICANT: PATEL, SATIL D.
; APPLICANT: MCARTHUR, JAMES G.
; TITLE OF INVENTION: ADENO-ASSOCIATED VIRUS CAPSID IMMUNOLOGIC DETERMINANTS
; FILE REFERENCE: 39672
; CURRENT APPLICATION NUMBER: US/10/293,478
; PRIOR FILING DATE: 2002-11-14
; PRIOR FILING DATE: 2002-11-14
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 1
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Adeno-associated virus
US-10-293-478-1

Query Match 100.0%; Score 3994; DB 14; Length 735;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAADGYLPDWLEDTLSGIRQWKLKPGPPPPPAERHKDDSRGLVLPKYKLGPNGLD 60
Db |||||
1 MAADGYLPDWLEDTLSGIRQWKLKPGPPPPPAERHKDDSRGLVLPKYKLGPNGLD 60
Qy 61 KGEPVNEADAAALEHDKAYDRQLDSGDNVYLYKNHDAEFQERLKEDTSFGNGLGRAVFQ 120
Db |||||
61 KGEPVNEADAAALEHDKAYDRQLDSGDNVYLYKNHDAEFQERLKEDTSFGNGLGRAVFQ 120

121 AKKRVLEPLGLVEEPVKTAPEGKRPVEHSPVPEPDSSTGKGAGQOPARKLNFGQTGDAD 180
Db |||||
121 AKKRVLEPLGLVEEPVKTAPEGKRPVEHSPVPEPDSSTGKGAGQOPARKLNFGQTGDAD 180
Qy |||||
Db |||||
181 SVDPDQPLGPPAAPSGLTNTMATGSGAPMADNNEGADGVGNSGNWHCDSTWMDRVI 240
Db |||||
181 SVDPDQPLGPPAAPSGLTNTMATGSGAPMADNNEGADGVGNSGNWHCDSTWMDRVI 240
Qy |||||
Db |||||
241 TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPWGYPDFNRFCHFSPRDWORLI 300
Db |||||
241 TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPWGYPDFNRFCHFSPRDWORLI 300
Qy |||||
Db |||||
301 NNNWFRPKRLNFKLNFQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQG 360
Db |||||
301 NNNWFRPKRLNFKLNFQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQG 360
Qy |||||
Db |||||
361 CLPPFPADVFWVPOQYGYLTLLNGSQAAGRSSFYCLEYFPQSOMLRTGNNFTSYTFEDVPF 420
Db |||||
361 CLPPFPADVFWVPOQYGYLTLLNGSQAAGRSSFYCLEYFPQSOMLRTGNNFTSYTFEDVPF 420
Qy |||||
Db |||||
421 HSSYAHQSQSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOFSQAQASDIRDQSRNMLPG 480
Db |||||
421 HSSYAHQSQSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOFSQAQASDIRDQSRNMLPG 480
Qy |||||
Db |||||
481 PCYRQORVSKTSADNNSEYSGWKATKYHLNGRDSLVPNGPAMASHKDDKEKFFPQSGVL 540
Db |||||
481 PCYRQORVSKTSADNNSEYSGWKATKYHLNGRDSLVPNGPAMASHKDDKEKFFPQSGVL 540
Qy |||||
Db |||||
541 IFKGQSEKTNVDIEKVMITDEEIRITNPVATEQYGSVSTNLQGRNQAATAADVNTQGV 600
Db |||||
541 IFKGQSEKTNVDIEKVMITDEEIRITNPVATEQYGSVSTNLQGRNQAATAADVNTQGV 600
Qy |||||
Db |||||
601 LPMWQORDVYLOQPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQIILIKNTVPANPSTT 660
Db |||||
601 LPMWQORDVYLOQPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQIILIKNTVPANPSTT 660
Qy |||||
Db |||||
661 FSAAKFASFIQYSTQGVSVIEWELOKENSKRNPEIQYTSNKNKSNVNDFTVDTNGVY 720
Db |||||
661 FSAAKFASFIQYSTQGVSVIEWELOKENSKRNPEIQYTSNKNKSNVNDFTVDTNGVY 720
Qy |||||
Db |||||
721 SEPRPIGTRYLTRNL 735
Db |||||
721 SEPRPIGTRYLTRNL 735

RESULT 3

US-10-291-583-70
; Sequence 70, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Virus (A
; FILE REFERENCE: UPN-02735USA
; CURRENT APPLICATION NUMBER: US/10/291,583
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US-60/350,607
; PRIOR FILING DATE: 2001-11-11
; PRIOR APPLICATION NUMBER: US 60/341,117
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,066
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,675
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 70
; LENGTH: 735
; TYPE: PRT
; ORGANISM: capsid protein of AAV serotype, clone AAV2

US-10-291-593-70

```
Query Match      100.0%; Score 3994; DB 14; Length 735;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADGYPDLWLEDTLSEGIQWKLKPGPPPPKPAERHKDDSRGLVLPYKYLGPFNGLD 60
DB 1 MAADGYPDLWLEDTLSEGIQWKLKPGPPPPKPAERHKDDSRGLVLPYKYLGPFNGLD 60

QY 61 KGEVNEADAALAHDKAYDRQLDSDGNPYLYKNHADAFAERLKHEDTSFGNIGRAVQ 120
DB 61 KGEVNEADAALAHDKAYDRQLDSDGNPYLYKNHADAFAERLKHEDTSFGNIGRAVQ 120

QY 121 AKKRVLEPLGLVEEVPKTAAGKRPVEHSPVEPDSSTGKAGQOQPAKRLNFGQTGDAD 180
DB 121 AKKRVLEPLGLVEEVPKTAAGKRPVEHSPVEPDSSTGKAGQOQPAKRLNFGQTGDAD 180

QY 181 SVDPDQPLGQPPAAPSGLTNTMATGSGAPMADNNEGADGVGNSGNWHCDSTWMDRVI 240
DB 181 SVDPDQPLGQPPAAPSGLTNTMATGSGAPMADNNEGADGVGNSGNWHCDSTWMDRVI 240

QY 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPMGYFDNRFCHFSRPRWQRLI 300
DB 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPMGYFDNRFCHFSRPRWQRLI 300

QY 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPVVLSAHQ 360
DB 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPVVLSAHQ 360

QY 361 CLPPPADVFMVPOGYLTLNNGSAQVGRSSFYCLEYFPSSQMLRTGNNFTFSYTFEDVVF 420
DB 361 CLPPPADVFMVPOGYLTLNNGSAQVGRSSFYCLEYFPSSQMLRTGNNFTFSYTFEDVVF 420

QY 421 HSSVAHSQSLDRLMNLIDQYLYLSRNTTSGTTSQSLQFSQAGASDIRQSRNWLPG 480
DB 421 HSSVAHSQSLDRLMNLIDQYLYLSRNTTSGTTSQSLQFSQAGASDIRQSRNWLPG 480

QY 481 PCYQORVSKTSADNNNSEYSGTATKYHLNGRDSLVPNGPAMASHKODEEKFPPQSGVL 540
DB 481 PCYQORVSKTSADNNNSEYSGTATKYHLNGRDSLVPNGPAMASHKODEEKFPPQSGVL 540

QY 541 IFGKQSEKTNVDIEKVMITDEEIRTTNPVATEQYGSVSTNLQRGNQAATADVNTQGV 600
DB 541 IFGKQSEKTNVDIEKVMITDEEIRTTNPVATEQYGSVSTNLQRGNQAATADVNTQGV 600

QY 601 LPGMWQDRDYYLQGPFWAKIPHTDGHFHPSPLMGGFGLKHPPPOILLIKNTVPANPSTT 660
DB 601 LPGMWQDRDYYLQGPFWAKIPHTDGHFHPSPLMGGFGLKHPPPOILLIKNTVPANPSTT 660

QY 661 FSAAKFASFIQYSTQGVSVLEWELQKNSKRWNPFIQYTSNKNKSVNVDFTVDTNGVY 720
DB 661 FSAAKFASFIQYSTQGVSVLEWELQKNSKRWNPFIQYTSNKNKSVNVDFTVDTNGVY 720
```

RESULT 4

```
US-10-423-704A-4
; Sequence 4, Application US/10423704A
; Publication No. US2003022828A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvira, Mauricio
; TITLE OF INVENTION: Adeno-Associated Virus (AAV) Serotype 8 Sequences, Vectors
; TITLE OF INVENTION: Containing Same, and Uses Therefor
; FILE REFERENCE: UPN-02733AUSA
; CURRENT APPLICATION NUMBER: US/10/423,704A
; CURRENT FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: US 60/341,151
```

RESULT 5

103

```
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US 60/377,133
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/386,122
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: PCT/US02/33630
; PRIOR FILING DATE: 2002-11-12
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 735
; TYPE: PRT
; ORGANISM: adeno-associated virus serotype 2
US-10-423-704A-4
```

```
Query Match      100.0%; Score 3994; DB 15; Length 735;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADGYPDLWLEDTLSEGIQWKLKPGPPPPKPAERHKDDSRGLVLPYKYLGPFNGLD 60
DB 1 MAADGYPDLWLEDTLSEGIQWKLKPGPPPPKPAERHKDDSRGLVLPYKYLGPFNGLD 60

QY 61 KGEVNEADAALAHDKAYDRQLDSDGNPYLYKNHADAFAERLKHEDTSFGNIGRAVQ 120
DB 61 KGEVNEADAALAHDKAYDRQLDSDGNPYLYKNHADAFAERLKHEDTSFGNIGRAVQ 120

QY 121 AKKRVLEPLGLVEEVPKTAAGKRPVEHSPVEPDSSTGKAGQOQPAKRLNFGQTGDAD 180
DB 121 AKKRVLEPLGLVEEVPKTAAGKRPVEHSPVEPDSSTGKAGQOQPAKRLNFGQTGDAD 180

QY 181 SVDPDQPLGQPPAAPSGLTNTMATGSGAPMADNNEGADGVGNSGNWHCDSTWMDRVI 240
DB 181 SVDPDQPLGQPPAAPSGLTNTMATGSGAPMADNNEGADGVGNSGNWHCDSTWMDRVI 240

QY 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPMGYFDNRFCHFSRPRWQRLI 300
DB 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPMGYFDNRFCHFSRPRWQRLI 300

QY 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPVVLSAHQ 360
DB 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPVVLSAHQ 360

QY 361 CLPPPADVFMVPOGYLTLNNGSAQVGRSSFYCLEYFPSSQMLRTGNNFTFSYTFEDVVF 420
DB 361 CLPPPADVFMVPOGYLTLNNGSAQVGRSSFYCLEYFPSSQMLRTGNNFTFSYTFEDVVF 420

QY 421 HSSVAHSQSLDRLMNLIDQYLYLSRNTTSGTTSQSLQFSQAGASDIRQSRNWLPG 480
DB 421 HSSVAHSQSLDRLMNLIDQYLYLSRNTTSGTTSQSLQFSQAGASDIRQSRNWLPG 480

QY 481 PCYQORVSKTSADNNNSEYSGTATKYHLNGRDSLVPNGPAMASHKODEEKFPPQSGVL 540
DB 481 PCYQORVSKTSADNNNSEYSGTATKYHLNGRDSLVPNGPAMASHKODEEKFPPQSGVL 540

QY 541 IFGKQSEKTNVDIEKVMITDEEIRTTNPVATEQYGSVSTNLQRGNQAATADVNTQGV 600
DB 541 IFGKQSEKTNVDIEKVMITDEEIRTTNPVATEQYGSVSTNLQRGNQAATADVNTQGV 600

QY 601 LPGMWQDRDYYLQGPFWAKIPHTDGHFHPSPLMGGFGLKHPPPOILLIKNTVPANPSTT 660
DB 601 LPGMWQDRDYYLQGPFWAKIPHTDGHFHPSPLMGGFGLKHPPPOILLIKNTVPANPSTT 660

QY 661 FSAAKFASFIQYSTQGVSVLEWELQKNSKRWNPFIQYTSNKNKSVNVDFTVDTNGVY 720
DB 661 FSAAKFASFIQYSTQGVSVLEWELQKNSKRWNPFIQYTSNKNKSVNVDFTVDTNGVY 720

QY 721 SEPRPIGTRYLTRNL 735
DB 721 SEPRPIGTRYLTRNL 735
```

```
US-10-496-799-4
; Sequence 4, Application US/10496799
; Publication No. US20050014262A1
; GENERAL INFORMATION:
; APPLICANT: The Trustees of The University of Pennsylvania
; APPLICANT: Geo, Guangping
; APPLICANT: Wilson, James M.
; APPLICANT: Alvirra, Mauricio
; TITLE OF INVENTION: Adeno-Associated Virus (AAV) Serotype 9 Sequences, Vectors Contai
; - FILE REFERENCE: UPN-027346CT
; CURRENT APPLICATION NUMBER: US/10/496,799
; PRIOR FILING DATE: 2004-06-08
; PRIOR FILING DATE: 2001-12-17
; PRIOR FILING DATE: 2001-12-17
; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 735
; TYPE: PRT
; ORGANISM: adeno-associated virus serotype 2
US-10-496-799-4

Query Match      100.0%; Score 3994; DB 17; Length 735;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADGVLDPDLEDTLSEGIQWKKLPGPPPPKPAERHKDDSRGLVPGYKYLGPFGNGLD 60
DB 1 MAADGVLDPDLEDTLSEGIQWKKLPGPPPPKPAERHKDDSRGLVPGYKYLGPFGNGLD 60
QY 61 KGEVNEADAAALEHDKAYDRQLDSGDNPLYKYNHADADEFQRLKEDTSFGNIGRAVFQ 120
DB 61 KGEVNEADAAALEHDKAYDRQLDSGDNPLYKYNHADADEFQRLKEDTSFGNIGRAVFQ 120
QY 121 AKKRVLEPLGLVEEPVKTA PGKRPVEHSPVEPDSSTGKAGQOPARKLNFGQTGDAD 180
DB 121 AKKRVLEPLGLVEEPVKTA PGKRPVEHSPVEPDSSTGKAGQOPARKLNFGQTGDAD 180
QY 181 SVDPDPLGQPPAAPSGLTNTMATGSGAPMADNNEGADVGNSSGNWHCDSTWMDRVI 240
DB 181 SVDPDPLGQPPAAPSGLTNTMATGSGAPMADNNEGADVGNSSGNWHCDSTWMDRVI 240
QY 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHVFGYSTPWGTFDNRHCHFSPRDWORLI 300
DB 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHVFGYSTPWGTFDNRHCHFSPRDWORLI 300
QY 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQ 360
DB 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQ 360
QY 361 CLPPPPADVPMVPOGYLTLLNGSQAVGRSSFCYCLEYFPSPQMLRTGNNTFTSYTFEDVFP 420
DB 361 CLPPPPADVPMVPOGYLTLLNGSQAVGRSSFCYCLEYFPSPQMLRTGNNTFTSYTFEDVFP 420
QY 421 HSSVAHSQSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDI RQSRNWLPG 480
DB 421 HSSVAHSQSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDI RQSRNWLPG 480
QY 481 PCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSL VNPFGPAMASHKDEKFFPQSGVL 540
DB 481 PCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSL VNPFGPAMASHKDEKFFPQSGVL 540
QY 541 IFGQSGSEKTNVDIEKVMITDEEIRTTNPVATPQYGSVSTNLQRGNQAATADVNTQGV 600
DB 541 IFGQSGSEKTNVDIEKVMITDEEIRTTNPVATPQYGSVSTNLQRGNQAATADVNTQGV 600
QY 601 LPMGMQDRDVLQGPITWAKI PHTDGHFSPLMGFGFLKHPPPQILIKNTVPANPSTT 660
DB 601 LPMGMQDRDVLQGPITWAKI PHTDGHFSPLMGFGFLKHPPPQILIKNTVPANPSTT 660
```

```
QY 661 FSAAKFASFTQYSTGQSVSEIWELOKENS KRWNPBIQYTSNYKSVNVDFVTDTNGVY 720
DB 661 FSAAKFASFTQYSTGQSVSEIWELOKENS KRWNPBIQYTSNYKSVNVDFVTDTNGVY 720
QY 721 SEPRPIGTRYLTRNL 735
DB 721 SEPRPIGTRYLTRNL 735
```

RESULT 6

```
US-10-880-297-11
; Sequence 11, Application US/10880297
; Publication No. US20050053922A1
; GENERAL INFORMATION:
; APPLICANT: SCHAFER, DAVID V.
; APPLICANT: KASPAR, BRIAN
; APPLICANT: MAESTRI, NARENDRA
; TITLE OF INVENTION: MUTANT ADENO-ASSOCIATED VIRUS VIRIONS
; FILE REFERENCE: BERK-024
; CURRENT APPLICATION NUMBER: US/10/880,297
; PRIOR FILING DATE: 2004-06-28
; PRIOR FILING DATE: 2003-06-30
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 735
; TYPE: PRT
; ORGANISM: adeno-associated virus-2
US-10-880-297-11
```

```
Query Match      99.9%; Score 3989; DB 17; Length 735;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 734; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 MAADGVLDPDLEDTLSEGIQWKKLPGPPPPKPAERHKDDSRGLVPGYKYLGPFGNGLD 60
DB 1 MAADGVLDPDLEDTLSEGIQWKKLPGPPPPKPAERHKDDSRGLVPGYKYLGPFGNGLD 60
QY 61 KGEVNEADAAALEHDKAYDRQLDSGDNPLYKYNHADADEFQRLKEDTSFGNIGRAVFQ 120
DB 61 KGEVNEADAAALEHDKAYDRQLDSGDNPLYKYNHADADEFQRLKEDTSFGNIGRAVFQ 120
QY 121 AKKRVLEPLGLVEEPVKTA PGKRPVEHSPVEPDSSTGKAGQOPARKLNFGQTGDAD 180
DB 121 AKKRVLEPLGLVEEPVKTA PGKRPVEHSPVEPDSSTGKAGQOPARKLNFGQTGDAD 180
QY 181 SVDPDPLGQPPAAPSGLTNTMATGSGAPMADNNEGADVGNSSGNWHCDSTWMDRVI 240
DB 181 SVDPDPLGQPPAAPSGLTNTMATGSGAPMADNNEGADVGNSSGNWHCDSTWMDRVI 240
QY 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHVFGYSTPWGTFDNRHCHFSPRDWORLI 300
DB 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHVFGYSTPWGTFDNRHCHFSPRDWORLI 300
QY 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQ 360
DB 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQ 360
QY 361 CLPPPPADVPMVPOGYLTLLNGSQAVGRSSFCYCLEYFPSPQMLRTGNNTFTSYTFEDVFP 420
DB 361 CLPPPPADVPMVPOGYLTLLNGSQAVGRSSFCYCLEYFPSPQMLRTGNNTFTSYTFEDVFP 420
QY 421 HSSVAHSQSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDI RQSRNWLPG 480
DB 421 HSSVAHSQSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDI RQSRNWLPG 480
QY 481 PCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSL VNPFGPAMASHKDEKFFPQSGVL 540
DB 481 PCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSL VNPFGPAMASHKDEKFFPQSGVL 540
QY 541 IFGQSGSEKTNVDIEKVMITDEEIRTTNPVATPQYGSVSTNLQRGNQAATADVNTQGV 600
```

```
Db 541 IFGKGSSEKTNVDIEKWMITDDEEIRTTNPVATEQYGSVSTNLQRGNQAATADVNTQGV 600
Qy 601 LPGMWQDRDYYLQGPPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILLIKNTVPANPSTT 660
Db 601 LPGMWQDRDYYLQGPPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILLIKNTVPANPSTT 660
Qy 661 FSAAKFASFITQYSTGQSVSEIWELOKENSKRWNPEIQYTSNYNKSNNVDFTVDTNGVY 720
Db 661 FSAAKFASFITQYSTGQSVSEIWELOKENSKRWNPEIQYTSNYNKSNNVDFTVDTNGVY 720
Qy 721 SEPRPIGTRYLTRNL 735
Db 721 SEPRPIGTRYLTRNL 735

RESULT 7
US-10-880-297-15
; Sequence 15, Application US/10880297
; Publication No. US20050053922A1
; GENERAL INFORMATION:
; APPLICANT: SCHAFER, DAVID V.
; APPLICANT: KASPAR, BRIAN
; APPLICANT: MAESTRI, NARENDRA
; TITLE OF INVENTION: MUTANT ADENO-ASSOCIATED VIRUS VIRIONS
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: BERK-024
; CURRENT APPLICATION NUMBER: US/10/880,297
; PRIOR FILING DATE: 2004-06-28
; PRIOR APPLICATION NUMBER: 60/484,111
; PRIOR FILING DATE: 2003-06-30
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 735
; TYPE: PRT
; ORGANISM: adeno-associated virus-2
US-10-880-297-15

Query Match 99.8%; Score 3985; DB 17; Length 735;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MAADGYLPDWLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPYKYLGPFGNGLD 60
Db 1 MAADGYLPDWLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPYKYLGPFGNGLD 60
Qy 61 KGEVNEADAAALEHDKAYDRQLDSDGNPNLYKYNHADAEPQERLKEDTSFGNLRGAVFQ 120
Db 61 KGEVNEADAAALEHDKAYDRQLDSDGNPNLYKYNHADAEPQERLKEDTSFGNLRGAVFQ 120
Qy 121 AKKRVLPLGLVEEPVKTAPGKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGOTGDAD 180
Db 121 AKKRVLPLGLVEEPVKTAPGKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGOTGDAD 180
Qy 181 SVDPDQPLGQPPAAPSGLTNTMATGSGAPMADNNEGADGVNSGNNHCDSTWMDRVI 240
Db 181 SVDPDQPLGQPPAAPSGLTNTMATGSGAPMADNNEGADGVNSGNNHCDSTWMDRVI 240
Qy 241 TTSTRWTALPTYNNHLYKQISSQSGASNDNHFGYSTPWGYPDFNRFCHFSPRDWRLLI 300
Db 241 TTSTRWTALPTYNNHLYKQISSQSGASNDNHFGYSTPWGYPDFNRFCHFSPRDWRLLI 300
Qy 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTTIANNLTSTVQVFTDSEYQLPVVLGSAHQG 360
Db 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTTIANNLTSTVQVFTDSEYQLPVVLGSAHQG 360
Qy 361 CLPPPPADVFWVPOGYLTLNNGSQAAGRSSFCYLEYFPQMLRTGNFTFSTYTFEDVFP 420
Db 361 CLPPPPADVFWVPOGYLTLNNGSQAAGRSSFCYLEYFPQMLRTGNFTFSTYTFEDVFP 420
Qy 421 HSSVAHSQSLDRLMNLPLIDQYLYYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
Db 421 HSSVAHSQSLDRLMNLPLIDQYLYYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
```

```
Db 421 HSSVAHSQSLDRLMNLPLIDQYLYYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
Qy 481 PCYRQORVSKTSADNNNSEYSWTGATKYHLNAGRDSLVPNPGPAMASHKDDDEKFFPQSGVYL 540
Db 481 PCYRQORVSKTSADNNNSEYSWTGATKYHLNAGRDSLVPNPGPAMASHKDDDEKFFPQSGVYL 540
Qy 541 IFGKGSSEKTNVDIEKWMITDDEEIRTTNPVATEQYGSVSTNLQRGNQAATADVNTQGV 600
Db 541 IFGKGSSEKTNVDIEKWMITDDEEIRTTNPVATEQYGSVSTNLQRGNQAATADVNTQGV 600
Qy 601 LPGMWQDRDYYLQGPPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILLIKNTVPANPSTT 660
Db 601 LPGMWQDRDYYLQGPPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILLIKNTVPANPSTT 660
Qy 661 FSAAKFASFITQYSTGQSVSEIWELOKENSKRWNPEIQYTSNYNKSNNVDFTVDTNGVY 720
Db 661 FSAAKFASFITQYSTGQSVSEIWELOKENSKRWNPEIQYTSNYNKSNNVDFTVDTNGVY 720
Qy 721 SEPRPIGTRYLTRNL 735
Db 721 SEPRPIGTRYLTRNL 735

RESULT 8
US-10-880-297-5
; Sequence 5, Application US/10880297
; Publication No. US20050053922A1
; GENERAL INFORMATION:
; APPLICANT: SCHAFER, DAVID V.
; APPLICANT: KASPAR, BRIAN
; APPLICANT: MAESTRI, NARENDRA
; TITLE OF INVENTION: MUTANT ADENO-ASSOCIATED VIRUS VIRIONS
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: BERK-024
; CURRENT APPLICATION NUMBER: US/10/880,297
; PRIOR FILING DATE: 2004-06-28
; PRIOR APPLICATION NUMBER: 60/484,111
; PRIOR FILING DATE: 2003-06-30
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 733
; TYPE: PRT
; ORGANISM: adeno-associated virus-2
US-10-880-297-5

Query Match 99.7%; Score 3984; DB 17; Length 733;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 733; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAADGYLPDWLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPYKYLGPFGNGLD 60
Db 1 MAADGYLPDWLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPYKYLGPFGNGLD 60
Qy 61 KGEVNEADAAALEHDKAYDRQLDSDGNPNLYKYNHADAEPQERLKEDTSFGNLRGAVFQ 120
Db 61 KGEVNEADAAALEHDKAYDRQLDSDGNPNLYKYNHADAEPQERLKEDTSFGNLRGAVFQ 120
Qy 121 AKKRVLPLGLVEEPVKTAPGKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGOTGDAD 180
Db 121 AKKRVLPLGLVEEPVKTAPGKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGOTGDAD 180
Qy 181 SVDPDQPLGQPPAAPSGLTNTMATGSGAPMADNNEGADGVNSGNNHCDSTWMDRVI 240
Db 181 SVDPDQPLGQPPAAPSGLTNTMATGSGAPMADNNEGADGVNSGNNHCDSTWMDRVI 240
Qy 241 TTSTRWTALPTYNNHLYKQISSQSGASNDNHFGYSTPWGYPDFNRFCHFSPRDWRLLI 300
Db 241 TTSTRWTALPTYNNHLYKQISSQSGASNDNHFGYSTPWGYPDFNRFCHFSPRDWRLLI 300
Qy 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTTIANNLTSTVQVFTDSEYQLPVVLGSAHQG 360
Db 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTTIANNLTSTVQVFTDSEYQLPVVLGSAHQG 360
```

```

QY 361 CLPPFPADVPMVPOYGYLTLLNGSQVGRSSFCYCLEYFSPQMLRTGNNFTSYTFEDVFF 420
Db 361 CLPPFPADVPMVPOYGYLTLLNGSQVGRSSFCYCLEYFSPQMLRTGNNFTSYTFEDVFF 420
QY 421 HSSYAHQSLSLRLMNLPLIDQYLYLSRTNTPSGTGTQSRLOPSQAGASDIRQSRNLPG 480
Db 421 HSSYAHQSLSLRLMNLPLIDQYLYLSRTNTPSGTGTQSRLOPSQAGASDIRQSRNLPG 480
QY 481 PCYRQORVSKTSADNNNSYSWTGATKYHLNGRDSLVNPGPAMASHKDDDEKFPQSGVL 540
Db 481 PCYRQORVSKTSADNNNSYSWTGATKYHLNGRDSLVNPGPAMASHKDDDEKFPQSGVL 540
QY 541 IFGQOSEKTNVDIEKWMITDDEEIRTTNPVATEQYGSVSTNLQRNQAAATADVNTQGV 600
Db 541 IFGQOSEKTNVDIEKWMITDDEEIRTTNPVATEQYGSVSTNLQRNQAAATADVNTQGV 600
QY 601 LFGWVWQDRDVLQGPIMAKIPIHTDGHFPHSPMLMGFGGLKHPPPQIILIKNTVPANPSTT 660
Db 601 LFGWVWQDRDVLQGPIMAKIPIHTDGHFPHSPMLMGFGGLKHPPPQIILIKNTVPANPSTT 660
QY 661 FSAAKFASFIQYSTQGVSVSEIWELOKENSQRWNPETIQTSTNYSKSVNVDFTVDTNGVY 720
Db 661 FSAAKFASFIQYSTQGVSVSEIWELOKENSQRWNPETIQTSTNYSKSVNVDFTVDTNGVY 720
QY 721 SEPRPIGTRYLTR 733
Db 721 SEPRPIGTRYLTR 733

```

RESULT 9

US-10-880-297-13

/ Sequence 13, Application US/10880297

/ Publication No. US20050053922A1

/ GENERAL INFORMATION:

/ APPLICANT: SCHAFER, DAVID V.

/ APPLICANT: KASPAR, BRIAN

/ APPLICANT: MAHESRI, NARENDRA

/ TITLE OF INVENTION: MUTANT ADENO-ASSOCIATED VIRUS VIRIONS

/ FILE REFERENCE: BERK-024

/ CURRENT APPLICATION NUMBER: US/10/880,297

/ PRIOR FILING DATE: 2004-06-28

/ PRIOR FILING DATE: 2003-06-30

/ NUMBER OF SEQ ID NOS: 41

/ SOFTWARE: FastSeq for Windows Version 4.0

/ SEQ ID NO 13

/ LENGTH: 735

/ TYPE: PRT

/ ORGANISM: adeno-associated virus-2

US-10-880-297-13

Query Match

Best Local Similarity 99.7%; Score 3984; DB 17; Length 735;

Matches 733; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAADGYLPDWLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPKYKLGPNGLD 60

Db 1 MAADGYLPDWLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPKYKLGPNGLD 60

QY 61 KGEVNEADAALAHDKAYDRQLDSDGNPYLYKNHADAFOERLKEDTSFGNLRGAVFQ 120

Db 61 KGEVNEADAALAHDKAYDRQLDSDGNPYLYKNHADAFOERLKEDTSFGNLRGAVFQ 120

QY 121 AKKRVLEPLGLVEPVKTAPEGKRPVESHSPVDPSSSGTGKAGQAPARKRNFQGTGDAD 180

Db 121 AKKRVLEPLGLVEPVKTAPEGKRPVESHSPVDPSSSGTGKAGQAPARKRNFQGTGDAD 180

QY 181 SVDPQPLGQPPAAPSGLTNTMATSGAPMADNEGADGVNSSGNWCHDSTWNGDRVI 240

Db 181 SVDPQPLGQPPAAPSGLTNTMATSGAPMADNEGADGVNSSGNWCHDSTWNGDRVI 240

```

QY 241 TTSTRTALPTYNHLYKQISSQSGASNDNHFCYSTPWCYFDPNRFCHFSRDMQRLI 300
Db 241 TTSTRTALPTYNHLYKQISSQSGASNDNHFCYSTPWCYFDPNRFCHFSRDMQRLI 300
QY 301 NNNWCFRPRKRLNFKLFIQVKEVTQNDGTTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQ 360
Db 301 NNNWCFRPRKRLNFKLFIQVKEVTQNDGTTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQ 360
QY 361 CLPPFPADVPMVPOYGYLTLLNGSQVGRSSFCYCLEYFSPQMLRTGNNFTSYTFEDVFF 420
Db 361 CLPPFPADVPMVPOYGYLTLLNGSQVGRSSFCYCLEYFSPQMLRTGNNFTSYTFEDVFF 420
QY 421 HSSYAHQSLSLRLMNLPLIDQYLYLSRTNTPSGTGTQSRLOPSQAGASDIRQSRNLPG 480
Db 421 HSSYAHQSLSLRLMNLPLIDQYLYLSRTNTPSGTGTQSRLOPSQAGASDIRQSRNLPG 480
QY 481 PCYRQORVSKTSADNNNSYSWTGATKYHLNGRDSLVNPGPAMASHKDDDEKFPQSGVL 540
Db 481 PCYRQORVSKTSADNNNSYSWTGATKYHLNGRDSLVNPGPAMASHKDDDEKFPQSGVL 540
QY 541 IFGQOSEKTNVDIEKWMITDDEEIRTTNPVATEQYGSVSTNLQRNQAAATADVNTQGV 600
Db 541 IFGQOSEKTNVDIEKWMITDDEEIRTTNPVATEQYGSVSTNLQRNQAAATADVNTQGV 600
QY 601 LFGWVWQDRDVLQGPIMAKIPIHTDGHFPHSPMLMGFGGLKHPPPQIILIKNTVPANPSTT 660
Db 601 LFGWVWQDRDVLQGPIMAKIPIHTDGHFPHSPMLMGFGGLKHPPPQIILIKNTVPANPSTT 660
QY 661 FSAAKFASFIQYSTQGVSVSEIWELOKENSQRWNPETIQTSTNYSKSVNVDFTVDTNGVY 720
Db 661 FSAAKFASFIQYSTQGVSVSEIWELOKENSQRWNPETIQTSTNYSKSVNVDFTVDTNGVY 720
QY 721 SEPRPIGTRYLTR 735
Db 721 SEPRPIGTRYLTR 735

```

RESULT 10

US-10-880-297-19

/ Sequence 19, Application US/10880297

/ Publication No. US20050053922A1

/ GENERAL INFORMATION:

/ APPLICANT: SCHAFER, DAVID V.

/ APPLICANT: KASPAR, BRIAN

/ APPLICANT: MAHESRI, NARENDRA

/ TITLE OF INVENTION: MUTANT ADENO-ASSOCIATED VIRUS VIRIONS

/ FILE REFERENCE: BERK-024

/ CURRENT APPLICATION NUMBER: US/10/880,297

/ PRIOR FILING DATE: 2004-06-28

/ PRIOR FILING DATE: 2003-06-30

/ NUMBER OF SEQ ID NOS: 41

/ SOFTWARE: FastSeq for Windows Version 4.0

/ SEQ ID NO 19

/ LENGTH: 735

/ TYPE: PRT

/ ORGANISM: adeno-associated virus-2

US-10-880-297-19

Query Match

Best Local Similarity 99.7%; Score 3984; DB 17; Length 735;

Matches 732; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAADGYLPDWLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPKYKLGPNGLD 60

Db 1 MAADGYLPDWLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPKYKLGPNGLD 60

QY 61 KGEVNEADAALAHDKAYDRQLDSDGNPYLYKNHADAFOERLKEDTSFGNLRGAVFQ 120

Db 61 KGEVNEADAALAHDKAYDRQLDSDGNPYLYKNHADAFOERLKEDTSFGNLRGAVFQ 120

QY 121 AKKRVLEPLGLVEPVKTAPEGKRPVESHSPVDPSSSGTGKAGQAPARKRNFQGTGDAD 180

```
Db 121 AKKRVLEPLGLVEEPVKTAPGKRPVEHSPVEPDSSTGKAGQOPARKRLNFGQTGDAD 180
Qy 181 SVDPDQPLGPPAAASGLGTNTMATGSGAPMADNNEGADVGNSSGNWHCDSTWMDRVI 240
Db 181 SVDPDQPLGPPAAASGLGTNTMATGSGAPMADNNEGADVGNSSGNWHCDSTWMDRVI 240
Qy 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPBGYDFNRFCHFSFPRDQRLI 300
Db 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPBGYDFNRFCHFSFPRDQRLI 300
Qy 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTIANNLTSTVQVFTDSEYQLPYVLGSAHQG 360
Db 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTIANNLTSTVQVFTDSEYQLPYVLGSAHQG 360
Qy 361 CLPPFPADVFWPQGYLTLNNGSOAVGRSSFCYCLEYFPSOMLRTGNNFTSYTFEDVPF 420
Db 361 CLPPFPADVFWPQGYLTLNNGSOAVGRSSFCYCLEYFPSOMLRTGNNFTSYTFEDVPF 420
Qy 421 HSSYAHQSLSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOPSQAGASDIRQSRNWLPG 480
Db 421 HSSYAHQSLSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOPSQAGASDIRQSRNWLPG 480
Qy 481 PCYRQORVSKTSADNNSEYSWTGATKYHLNGRDSLVNPGPAMASHKDDKEKFFPQSGVL 540
Db 481 PCYRQORVSKTSADNNSEYSWTGATKYHLNGRDSLVNPGPAMASHKDDKEKFFPQSGVL 540
Qy 541 IFGKGSEKTNVDIEKWMITDEEERTTNPVATEQVGSVSTNLQGRNQAAATADVNTQGV 600
Db 541 IFGKGSEKTNVDIEKWMITDEEERTTNPVATEQVGSVSTNLQGRNQAAATADVNTQGV 600
Qy 601 LPMWQVQDRDVLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
Db 601 LPMWQVQDRDVLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
Qy 661 FSAAKFASFITQYSTGQSVSEIWELOKENSKRWNPEIQYTSNKNYSVNVDFTVDTNGVY 720
Db 661 FSAAKFASFITQYSTGQSVSEIWELOKENSKRWNPEIQYTSNKNYSVNVDFTVDTNGVY 720
Qy 721 SEPRPIGTRYLTRNL 735
Db 721 SEPRPIGTRYLTRNL 735
```

RESULT 11

US-10-880-297-7

```
; Sequence 7, Application US/10880297
; Publication No. US20050053922A1
; GENERAL INFORMATION:
; APPLICANT: SCHAFER, DAVID V.
; APPLICANT: KASPAR, BRIAN
; APPLICANT: MAHESRI, NARENDRA
; TITLE OF INVENTION: MUTANT ADENO-ASSOCIATED VIRUS VIRIONS
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: BERK-024
; CURRENT APPLICATION NUMBER: US/10/880,297
; PRIOR FILING DATE: 2004-06-28
; PRIOR APPLICATION NUMBER: 60/484,111
; PRIOR FILING DATE: 2003-06-30
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 735
; TYPE: PRT
; ORGANISM: adeno-associated virus-2
US-10-880-297-7
```

```
Query Match 99.7%; Score 3983; DB 17; Length 735;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 733; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MAADGYLPDWLEDTLSEGIQWKKLPKPPPPKPAERHKDDSRGLVLPYKYLGPFNGLD 60
|||||
```

```
Db 1 MAADGYLPDWLEDTLSEGIQWKKLPKPPPPKPAERHKDDSRGLVLPYKYLGPFNGLD 60
Qy 61 KGEVNVNEADAAAEHDKAYDROLDSGDNPLKYNHADADEFQERLKEDTSFGNGLGRAVQ 120
Db 61 KGEVNVNEADAAAEHDKAYDROLDSGDNPLKYNHADADEFQERLKEDTSFGNGLGRAVQ 120
Qy 121 AKKRVLEPLGLVEEPVKTAPGKRPVEHSPVEPDSSTGKAGQOPARKRLNFGQTGDAD 180
Db 121 AKKRVLEPLGLVEEPVKTAPGKRPVEHSPVEPDSSTGKAGQOPARKRLNFGQTGDAD 180
Qy 181 SVDPDQPLGPPAAASGLGTNTMATGSGAPMADNNEGADVGNSSGNWHCDSTWMDRVI 240
Db 181 SVDPDQPLGPPAAASGLGTNTMATGSGAPMADNNEGADVGNSSGNWHCDSTWMDRVI 240
Qy 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPBGYDFNRFCHFSFPRDQRLI 300
Db 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPBGYDFNRFCHFSFPRDQRLI 300
Qy 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTIANNLTSTVQVFTDSEYQLPYVLGSAHQG 360
Db 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTIANNLTSTVQVFTDSEYQLPYVLGSAHQG 360
Qy 361 CLPPFPADVFWPQGYLTLNNGSOAVGRSSFCYCLEYFPSOMLRTGNNFTSYTFEDVPF 420
Db 361 CLPPFPADVFWPQGYLTLNNGSOAVGRSSFCYCLEYFPSOMLRTGNNFTSYTFEDVPF 420
Qy 421 HSSYAHQSLSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOPSQAGASDIRQSRNWLPG 480
Db 421 HSSYAHQSLSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOPSQAGASDIRQSRNWLPG 480
Qy 481 PCYRQORVSKTSADNNSEYSWTGATKYHLNGRDSLVNPGPAMASHKDDKEKFFPQSGVL 540
Db 481 PCYRQORVSKTSADNNSEYSWTGATKYHLNGRDSLVNPGPAMASHKDDKEKFFPQSGVL 540
Qy 541 IFGKGSEKTNVDIEKWMITDEEERTTNPVATEQVGSVSTNLQGRNQAAATADVNTQGV 600
Db 541 IFGKGSEKTNVDIEKWMITDEEERTTNPVATEQVGSVSTNLQGRNQAAATADVNTQGV 600
Qy 601 LPMWQVQDRDVLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
Db 601 LPMWQVQDRDVLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
Qy 661 FSAAKFASFITQYSTGQSVSEIWELOKENSKRWNPEIQYTSNKNYSVNVDFTVDTNGVY 720
Db 661 FSAAKFASFITQYSTGQSVSEIWELOKENSKRWNPEIQYTSNKNYSVNVDFTVDTNGVY 720
Qy 721 SEPRPIGTRYLTRNL 735
Db 721 SEPRPIGTRYLTRNL 735
```

RESULT 12

```
US-10-880-297-9
; Sequence 9, Application US/10880297
; Publication No. US20050053922A1
; GENERAL INFORMATION:
; APPLICANT: SCHAFER, DAVID V.
; APPLICANT: KASPAR, BRIAN
; APPLICANT: MAHESRI, NARENDRA
; TITLE OF INVENTION: MUTANT ADENO-ASSOCIATED VIRUS VIRIONS
; TITLE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: BERK-024
; CURRENT APPLICATION NUMBER: US/10/880,297
; CURRENT FILING DATE: 2004-06-28
; PRIOR FILING DATE: 2003-06-30
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 735
; TYPE: PRT
; ORGANISM: adeno-associated virus-2
US-10-880-297-9
```

```
Query Match      99.6%; Score 3980; DB 17; Length 735;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 732; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy. 1 MAADGYLPDLEDTLSEGIQWMLKPGPPPPPAERHKDDSGRLVLPYKYLGPNGLD 60
Db 1 MAADGYLPDLEDTLSEGIQWMLKPGPPPPPAERHKDDSGRLVLPYKYLGPNGLD 60

Qy 61 KGEFVNEADAAALEHDKAYDRQLDSDGNPYLYKNHADAEFQERLKEDTSFGNLRGAVFQ 120
Db 61 KGEFVNEADAAALEHDKAYDRQLDSDGNPYLYKNHADAEFQERLKEDTSFGNLRGAVFQ 120

Qy 121 AKKRVLEPLGLVEEPVKTAPGKKRPVEHSPVDPSSGSGTGKAGQQPARKLNFQGTGDAD 180
Db 121 AKKRVLEPLGLVEEPVKTAPGKKRPVEHSPVDPSSGSGTGKAGQQPARKLNFQGTGDAD 180

Qy 181 SVDPDQPLGPPAAPSGLGTNTMATSGAPMADNNEGADGVNSGNWHCDSTWMDRVI 240
Db 181 SVDPDQPLGPPAAPSGLGTNTMATSGAPMADNNEGADGVNSGNWHCDSTWMDRVI 240

Qy 241 TTSRTWALPTNNHLYKQISSQSGASNDNHYPGISTPMGYDPDFNRHCHFSRDPWQRLI 300
Db 241 TTSRTWALPTNNHLYKQISSQSGASNDNHYPGISTPMGYDPDFNRHCHFSRDPWQRLI 300

Qy 301 NNNWGPRLKFLNFIQVKEVTQNDGTTTIANLNTSTVQVFTDSEYQLPVLGSAHQ 360
Db 301 NNNWGPRLKFLNFIQVKEVTQNDGTTTIANLNTSTVQVFTDSEYQLPVLGSAHQ 360

Qy 361 CLPPPADVFMVPOGYLTNNGSAQVGRSSPYCLEYFPSSQMLRTGNNFTSYTFEDVPF 420
Db 361 CLPPPADVFMVPOGYLTNNGSAQVGRSSPYCLEYFPSSQMLRTGNNFTSYTFEDLPF 420

Qy 421 HSSYAHQSGLRLMPLDQVLYLSRTNTPSGTTTQSRLOFSQAGSADIRDOGRNWLPG 480
Db 421 HSSYAHQSGLRLMPLDQVLYLSRTNTPSGTTTQSRLOFSQAGSADIRDOGRNWLPG 480

Qy 481 PCYRQORVSKTSADNNSEYSGWTGATKYHLNGRDSLVPNGPAMASHKDDKFFPQSGVL 540
Db 481 PCYRQORVSKTSADNNSEYSGWTGATKYHLNGRDSLVPNGPAMASHKDDKFFPQSGVL 540

Qy 541 IFGKQSEKTNVDIEKVMITDEEIRTNTPVATQYGSVSTNLRGNRQAATADVNTQGV 600
Db 541 IFGKQSEKTNVDIEKVMITDEEIRTNTPVATQYGSVSTNLRGNRQAATADVNTQGV 600

Qy 601 LFGMWQDRDYYLQSPIWAKIPHTDGHFHPSPLMGGFGLKHPPIQLIKNTPVPANPSTT 660
Db 601 LFGMWQDRDYYLQSPIWAKIPHTDGHFHPSPLMGGFGLKHPPIQLIKNTPVPANPSTT 660

Qy 661 FSAAKFASFIQYSTGVSVIEIWELOKENSKRWNPEIQYTSNKNKSVNVDFTVDNGVY 720
Db 661 FSAAKFASFIQYSTGVSVIEIWELOKENSKRWNPEIQYTSNKNKSVNVDFTVDNGVY 720

Qy 721 SEPRPIGTRYLTRNL 735
Db 721 SEPRPIGTRYLTRNL 735

RESULT 14
US-10-880-297-31
; Sequence 31, Application US/10880297
; Publication No. US20050053922A1
; GENERAL INFORMATION:
; APPLICANT: SCHAFER, DAVID V.
; APPLICANT: KASPAR, BRIAN
; APPLICANT: MAESTRI, NARENDRA
; TITLE OF INVENTION: MUTANT ADENO-ASSOCIATED VIRUS VIRIONS
; FILE OF INVENTION: AND METHODS OF USE THEREOF
; FILE REFERENCE: BERK-024
; CURRENT APPLICATION NUMBER: US/10/880, 297
; PRIOR APPLICATION NUMBER: 2004-06-28
; PRIOR APPLICATION NUMBER: 60/484,111
; PRIOR FILING DATE: 2003-06-30

; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: RastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 735
; TYPE: PRT
; ORGANISM: adeno-associated virus-2
US-10-880-297-31

Query Match      99.6%; Score 3980; DB 17; Length 735;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MAADGYLPDLEDTLSEGIQWMLKPGPPPPPAERHKDDSGRLVLPYKYLGPNGLD 60
Db 1 MAADGYLPDLEDTLSEGIQWMLKPGPPPPPAERHKDDSGRLVLPYKYLGPNGLD 60

Qy 61 KGEFVNEADAAALEHDKAYDRQLDSDGNPYLYKNHADAEFQERLKEDTSFGNLRGAVFQ 120
Db 61 KGEFVNEADAAALEHDKAYDRQLDSDGNPYLYKNHADAEFQERLKEDTSFGNLRGAVFQ 120

Qy 121 AKKRVLEPLGLVEEPVKTAPGKKRPVEHSPVDPSSGSGTGKAGQQPARKLNFQGTGDAD 180
Db 121 AKKRVLEPLGLVEEPVKTAPGKKRPVEHSPVDPSSGSGTGKAGQQPARKLNFQGTGDAD 180

Qy 181 SVDPDQPLGPPAAPSGLGTNTMATSGAPMADNNEGADGVNSGNWHCDSTWMDRVI 240
Db 181 SVDPDQPLGPPAAPSGLGTNTMATSGAPMADNNEGADGVNSGNWHCDSTWMDRVI 240

Qy 241 TTSRTWALPTNNHLYKQISSQSGASNDNHYPGISTPMGYDPDFNRHCHFSRDPWQRLI 300
Db 241 TTSRTWALPTNNHLYKQISSQSGASNDNHYPGISTPMGYDPDFNRHCHFSRDPWQRLI 300

Qy 301 NNNWGPRLKFLNFIQVKEVTQNDGTTTIANLNTSTVQVFTDSEYQLPVLGSAHQ 360
Db 301 NNNWGPRLKFLNFIQVKEVTQNDGTTTIANLNTSTVQVFTDSEYQLPVLGSAHQ 360

Qy 361 CLPPPADVFMVPOGYLTNNGSAQVGRSSPYCLEYFPSSQMLRTGNNFTSYTFEDVPF 420
Db 361 CLPPPADVFMVPOGYLTNNGSAQVGRSSPYCLEYFPSSQMLRTGNNFTSYTFEDLPF 420

Qy 421 HSSYAHQSGLRLMPLDQVLYLSRTNTPSGTTTQSRLOFSQAGSADIRDOGRNWLPG 480
Db 421 HSSYAHQSGLRLMPLDQVLYLSRTNTPSGTTTQSRLOFSQAGSADIRDOGRNWLPG 480

Qy 481 PCYRQORVSKTSADNNSEYSGWTGATKYHLNGRDSLVPNGPAMASHKDDKFFPQSGVL 540
Db 481 PCYRQORVSKTSADNNSEYSGWTGATKYHLNGRDSLVPNGPAMASHKDDKFFPQSGVL 540

Qy 541 IFGKQSEKTNVDIEKVMITDEEIRTNTPVATQYGSVSTNLRGNRQAATADVNTQGV 600
Db 541 IFGKQSEKTNVDIEKVMITDEEIRTNTPVATQYGSVSTNLRGNRQAATADVNTQGV 600

Qy 601 LFGMWQDRDYYLQSPIWAKIPHTDGHFHPSPLMGGFGLKHPPIQLIKNTPVPANPSTT 660
Db 601 LFGMWQDRDYYLQSPIWAKIPHTDGHFHPSPLMGGFGLKHPPIQLIKNTPVPANPSTT 660

Qy 661 FSAAKFASFIQYSTGVSVIEIWELOKENSKRWNPEIQYTSNKNKSVNVDFTVDNGVY 720
Db 661 FSAAKFASFIQYSTGVSVIEIWELOKENSKRWNPEIQYTSNKNKSVNVDFTVDNGVY 720

Qy 721 SEPRPIGTRYLTRNL 735
Db 721 SEPRPIGTRYLTRNL 735
```


;
; TITLE OF INVENTION: MUTANT ADENO-ASSOCIATED VIRUS VIRIONS
; FILE REFERENCE: BERK-024
; CURRENT APPLICATION NUMBER: US/10/880,297
; CURRENT FILING DATE: 2004-06-28
; PRIOR APPLICATION NUMBER: 60/484,111
; PRIOR FILING DATE: 2003-06-30
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 735
; TYPE: PRT
; ORGANISM: adeno-associated virus-2
US-10-880-297-25

Query Match 99.5%; Score 3975; DB 17; Length 735;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 731; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MAADGYLPDWLEDTLSEGIQWKLKPGPPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD 60
Db 1 MAADGYLPDWLEDTLSEGIQWKLKPGPPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD 60
Qy 61 KGEVNEADAALAHDKAYDRQLDSGDNPYLYKYNHADAFAEFQERLKEDTSFGNLRGAVFQ 120
Db 61 KGEVNEADAALAHDKAYDRQLDSGDNPYLYKYNHADAFAEFQERLKEDTSFGNLRGAVFQ 120
Qy 121 AKKRVLEPLGLVEEVPKTPAGKKRPVEHSPVEPDSSSGTGKAGQOPARKRLNFGOTGDAD 180
Db 121 AKKRVLEPLGLVEEVPKTPAGKKRPVEHSPVEPDSSSGTGKAGQOPARKRLNFGOTGDAD 180
Qy 181 SVDPDQPLGQPPAAPSGLTNTMTATGSGAPMADNNEGADGVNNSGNHCDSTWMDRVI 240
Db 181 SVDPDQPLGQPPAAPSGLTNTMTATGSGAPMADNNEGADGVNNSGNHCDSTWMDRVI 240
Qy 241 TTSTRTWALPTYNNHLYKQISSQSGASNDNHFGYSTPWGYPDFNRFCHFSRDPWQRLI 300
Db 241 TTSTRTWALPTYNNHLYKQISSQSGASNDNHFGYSTPWGYPDFNRFCHFSRDPWQRLI 300
Qy 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPVYLSAHQ 360
Db 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPVYLSAHQ 360
Qy 361 CLPPFPADVFPVQGYLTLLNGSOAVGRSSFYCLEYFPSSQMLRTGNNFTSYTFEDVPF 420
Db 361 CLPPFPADVFPVQGYLTLLNGSOAVGRSSFYCLEYFPSSQMLRTGNNFTSYTFEDVPF 420
Qy 421 HSSVAHSQSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
Db 421 HSSVAHSQSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
Qy 481 PCYQQRVSKTSADNNSEYSWTGATKYLHNGRDSLVPNGPAMASHKDDDEKFPQSGVL 540
Db 481 PCYQQRVSKTSADNNSEYSWTGATKYLHNGRDSLVPNGPAMASHKDDDEKFPQSGVL 540
Qy 541 IFGQSGSEKTNVDIEKMITDEEIRTNVATEQYGSVSTNLQRGNAQATADVNTQGV 600
Db 541 IFGQSGSEKTNVDIEKMITDEEIRTNVATEQYGSVSTNLQRGNAQATADVNTQGV 600
Qy 601 LPGMWQDRDYLQGPVWAKIPHTDGHFHPSPMLGGFGLKHPPPQIILKNTVPVNPSTT 660
Db 601 LPGMWQDRDYLQGPVWAKIPHTDGHFHPSPMLGGFGLKHPPPQIILKNTVPVNPSTT 660
Qy 661 FSAAKFASFITQYSTGQSVSEIWELOKENSKRWNPEIQYTSNNYKSNVNVFTVDTNGVY 720
Db 661 FSAAKFASFITQYSTGQSVSEIWELOKENSKRWNPEIQYTSNNYKSNVNVFTVDTNGVY 720
Qy 721 SEPRPIGTRYLTRNL 735
Db 721 SEPRPIGTRYLTRNL 735

US-10-880-297-27
; Sequence 27, Application US/10880297
; Publication No. US20050053922A1
; GENERAL INFORMATION:
; APPLICANT: SCHAFER, DAVID V.
; APPLICANT: KASPAR, BRIAN
; APPLICANT: MAHESTRY, NARENDRA
; TITLE OF INVENTION: MUTANT ADENO-ASSOCIATED VIRUS VIRIONS
; FILE REFERENCE: BERK-024
; CURRENT APPLICATION NUMBER: US/10/880,297
; CURRENT FILING DATE: 2004-06-28
; PRIOR APPLICATION NUMBER: 60/484,111
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 735
; TYPE: PRT
; ORGANISM: adeno-associated virus-2
US-10-880-297-27

Query Match 99.5%; Score 3975; DB 17; Length 735;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 732; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 MAADGYLPDWLEDTLSEGIQWKLKPGPPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD 60
Db 1 MAADGYLPDWLEDTLSEGIQWKLKPGPPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD 60
Qy 61 KGEVNEADAALAHDKAYDRQLDSGDNPYLYKYNHADAFAEFQERLKEDTSFGNLRGAVFQ 120
Db 61 KGEVNEADAALAHDKAYDRQLDSGDNPYLYKYNHADAFAEFQERLKEDTSFGNLRGAVFQ 120
Qy 121 AKKRVLEPLGLVEEVPKTPAGKKRPVEHSPVEPDSSSGTGKAGQOPARKRLNFGOTGDAD 180
Db 121 AKKRVLEPLGLVEEVPKTPAGKKRPVEHSPVEPDSSSGTGKAGQOPARKRLNFGOTGDAD 180
Qy 181 SVDPDQPLGQPPAAPSGLTNTMTATGSGAPMADNNEGADGVNNSGNHCDSTWMDRVI 240
Db 181 SVDPDQPLGQPPAAPSGLTNTMTATGSGAPMADNNEGADGVNNSGNHCDSTWMDRVI 240
Qy 241 TTSTRTWALPTYNNHLYKQISSQSGASNDNHFGYSTPWGYPDFNRFCHFSRDPWQRLI 300
Db 241 TTSTRTWALPTYNNHLYKQISSQSGASNDNHFGYSTPWGYPDFNRFCHFSRDPWQRLI 300
Qy 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPVYLSAHQ 360
Db 301 NNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPVYLSAHQ 360
Qy 361 CLPPFPADVFPVQGYLTLLNGSOAVGRSSFYCLEYFPSSQMLRTGNNFTSYTFEDVPF 420
Db 361 CLPPFPADVFPVQGYLTLLNGSOAVGRSSFYCLEYFPSSQMLRTGNNFTSYTFEDVPF 420
Qy 421 HSSVAHSQSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
Db 421 HSSVAHSQSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
Qy 481 PCYQQRVSKTSADNNSEYSWTGATKYLHNGRDSLVPNGPAMASHKDDDEKFPQSGVL 540
Db 481 PCYQQRVSKTSADNNSEYSWTGATKYLHNGRDSLVPNGPAMASHKDDDEKFPQSGVL 540
Qy 541 IFGQSGSEKTNVDIEKMITDEEIRTNVATEQYGSVSTNLQRGNAQATADVNTQGV 600
Db 541 IFGQSGSEKTNVDIEKMITDEEIRTNVATEQYGSVSTNLQRGNAQATADVNTQGV 600
Qy 601 LPGMWQDRDYLQGPVWAKIPHTDGHFHPSPMLGGFGLKHPPPQIILKNTVPVNPSTT 660
Db 601 LPGMWQDRDYLQGPVWAKIPHTDGHFHPSPMLGGFGLKHPPPQIILKNTVPVNPSTT 660
Qy 661 FSAAKFASFITQYSTGQSVSEIWELOKENSKRWNPEIQYTSNNYKSNVNVFTVDTNGVY 720
Db 661 FSAAKFASFITQYSTGQSVSEIWELOKENSKRWNPEIQYTSNNYKSNVNVFTVDTNGVY 720

QY 721 SEPRIGTRYLRNL 735
Db |||||
721 SEPRIGTRYLRNL 735

Search completed: May 19, 2005, 13:46:05
Job time : 142 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 13:24:55 ; Search time 42 Seconds
(without alignments)
1306.359 Million cell updates/sec

Title: US-10-038-972A-13
Perfect score: 3994
Sequence: 1 MAADGYLPDWLEDTLSEGI.....TNGVSEPRPIGTRYLTRL 735

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA*
1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCITUS-COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3994	100.0	735	US-09-321-589-1	Sequence 1, Appli
2	3994	100.0	735	US-10-293-478-1	Sequence 1, Appli
3	3402.5	85.2	736	US-09-807-802A-3	Sequence 3, Appli
4	3402.5	85.2	736	US-09-807-802A-13	Sequence 13, Appli
5	2764.5	69.2	599	US-09-807-802A-15	Sequence 15, Appli
6	2481.5	62.1	534	US-09-807-802A-17	Sequence 17, Appli
7	2397	60.0	734	US-09-532-594B-4	Sequence 4, Appli
8	1797	45.0	598	US-09-532-594B-16	Sequence 16, Appli
9	1669.5	41.8	544	US-09-532-594B-18	Sequence 18, Appli
10	1579.5	39.5	756	US-09-438-268-4	Sequence 4, Appli
11	486	12.2	543	US-08-856-841-22	Sequence 22, Appli
12	431	10.8	500	US-08-856-841-16	Sequence 16, Appli
13	431	10.8	501	US-08-856-841-18	Sequence 18, Appli
14	428	10.7	486	US-08-856-841-19	Sequence 19, Appli
15	327	8.2	415	US-08-856-841-20	Sequence 20, Appli
16	321	8.0	395	US-08-856-841-13	Sequence 13, Appli
17	312	7.8	264	US-08-856-841-14	Sequence 14, Appli
18	308	7.7	398	US-08-856-841-21	Sequence 21, Appli
19	282.5	7.1	387	US-08-856-841-17	Sequence 17, Appli
20	241.5	6.0	579	5223424-13	Patent No. 5223424
21	241.5	6.0	579	5223424-13	Patent No. 5223424
22	208	5.2	584	US-09-022-949-2	Sequence 2, Appli
23	138	3.5	227	US-08-856-841-15	Sequence 15, Appli
24	138	3.5	250	US-08-856-841-12	Sequence 12, Appli
25	136	3.4	210	US-08-856-841-9	Sequence 9, Appli
26	119	3.0	655	US-08-469-202-27	Sequence 27, Appli
27	119	3.0	655	US-08-484-434C-34	Sequence 34, Appli

28	119	3.0	655	4	US-09-384-361-34	Sequence 34, Appli
29	118	3.0	1095	4	US-09-107-532A-3855	Sequence 3855, Ap
30	117	2.9	560	3	US-08-814-052-6	Sequence 6, Appli
31	117	2.9	560	3	US-08-812-829-6	Sequence 6, Appli
32	117	2.9	847	4	US-09-373-157-4	Sequence 4, Appli
33	117	2.9	875	4	US-09-949-016-8582	Sequence 8582, Ap
34	117	2.9	1742	4	US-09-386-962C-4	Sequence 4, Appli
35	117	2.9	1742	4	US-09-386-959-4	Sequence 4, Appli
36	115.5	2.9	571	3	US-09-134-001C-3865	Sequence 3865, Ap
37	115	2.9	1155	4	US-09-710-279-1780	Sequence 1780, Ap
38	114	2.9	361	3	US-08-874-569B-21	Sequence 21, Appli
39	114	2.9	361	3	US-09-955-518-21	Sequence 21, Appli
40	113.5	2.8	501	4	US-09-248-796A-15119	Sequence 15119, A
41	113	2.8	983	4	US-09-538-092-1320	Sequence 1320, Ap
42	112	2.8	617	4	US-09-248-796A-26692	Sequence 26692, A
43	111.5	2.8	845	4	US-09-328-352-7017	Sequence 7017, Ap
44	111	2.8	624	3	US-08-947-965-78	Sequence 78, Appli
45	111	2.8	655	1	US-08-469-202-28	Sequence 28, Appli

ALIGNMENTS

RESULT 1
US-09-321-589-1
; Sequence 1, Application US/09321589
; Patent No. 6498244
; GENERAL INFORMATION:
; APPLICANT: PATEL, SALIL D.
; APPLICANT: MCARTHUR, JAMES G.
; TITLE OF INVENTION: ADENO-ASSOCIATED VIRUS CAPSID IMMUNOLOGIC DETERMINANTS
; FILE REFERENCE: 39672
; CURRENT APPLICATION NUMBER: US/09/321.589
; CURRENT FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 1
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Adeno-associated virus
US-09-321-589-1

Query Match 100.0%; Score 3994; DB 4; Length 735;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAADGYLPDWLEDTLSEGI	ROWWKLPKPPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD	60
Db	1	MAADGYLPDWLEDTLSEGI	ROWWKLPKPPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLD	60
Qy	61	KGEPVNEADAAALEHDKAYDQ	LDSDGNPNYLYKNHADAEFOERLKEDTSFGNIGRAVFO	120
Db	61	KGEPVNEADAAALEHDKAYDQ	LDSDGNPNYLYKNHADAEFOERLKEDTSFGNIGRAVFO	120
Qy	121	AKKRVLEPLGLVEEPVKTA	PGKRPVEHSVPEDSSSGTGAKGQQQPARKLNFGTGDAD	180
Db	121	AKKRVLEPLGLVEEPVKTA	PGKRPVEHSVPEDSSSGTGAKGQQQPARKLNFGTGDAD	180
Qy	181	SVPDPQIPGQPPAASGLG	TMTATGSCAPMADNNEGADGVNSGNHCDSTWMDRVI	240
Db	181	SVPDPQIPGQPPAASGLG	TMTATGSCAPMADNNEGADGVNSGNHCDSTWMDRVI	240
Qy	241	TTSTRFTWALPYNNHLYKQ	ISSQSGASNDNHFGYSTPWGYPDFNRFCHFSPRDQRLI	300
Db	241	TTSTRFTWALPYNNHLYKQ	ISSQSGASNDNHFGYSTPWGYPDFNRFCHFSPRDQRLI	300
Qy	301	NNNGFRPKRLNFKLFNTQ	VKVEVTQNDCTTTIANNLTSVQVFTDSEYQLPYVLGSAHQ	360
Db	301	NNNGFRPKRLNFKLFNTQ	VKVEVTQNDCTTTIANNLTSVQVFTDSEYQLPYVLGSAHQ	360
Qy	361	CLPPPPADVFVMPQGYLT	LNNGSAQGRSSFCYLEYFPQMLRTGNNFTFSTYTFEDVPF	420
Db	361	CLPPPPADVFVMPQGYLT	LNNGSAQGRSSFCYLEYFPQMLRTGNNFTFSTYTFEDVPF	420

QY 421 HSSVAHSOSLRLANPLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
DB 421 HSSVAHSOSLRLANPLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
QY 481 PCYRQORVSKTSADNNSEYSGWTGATKYHLNGRDSLVPNPGPAMASHKDDKEKFFPQSGVL 540
DB 481 PCYRQORVSKTSADNNSEYSGWTGATKYHLNGRDSLVPNPGPAMASHKDDKEKFFPQSGVL 540
QY 541 IFGKGSEKTNVDIEKWMITDEBEIRTNPNVATEQYGSVSTNLQRNQAATAADVNTQGV 600
DB 541 IFGKGSEKTNVDIEKWMITDEBEIRTNPNVATEQYGSVSTNLQRNQAATAADVNTQGV 600
QY 601 LPGVMWQDRDVLQGPWAKIPHTDGHFHPSPLMGGFGLKHPHPQILIKNTVPVNPSTT 660
DB 601 LPGVMWQDRDVLQGPWAKIPHTDGHFHPSPLMGGFGLKHPHPQILIKNTVPVNPSTT 660
QY 661 FSAAKFASFTQYSTQGVSVSEIWELOKENSKEWNPFIQYTSNYSKSVNVDFTVDTNGVY 720
DB 661 FSAAKFASFTQYSTQGVSVSEIWELOKENSKEWNPFIQYTSNYSKSVNVDFTVDTNGVY 720
QY 721 SEPRPIGTRYLTRNL 735
DB 721 SEPRPIGTRYLTRNL 735

RESULT 2

US-10-293-478-1
; Sequence 1, Application US/10293478
; Patent No. 6733757
; GENERAL INFORMATION:
; APPLICANT: PATEL, SALIL D.
; APPLICANT: MCARTHUR, JAMES G.
; TITLE OF INVENTION: ADENO-ASSOCIATED VIRUS CAPSID IMMUNOLOGIC DETERMINANTS
; FILE REFERENCE: 39672
; CURRENT APPLICATION NUMBER: US/10/293,478
; PRIOR FILING DATE: 2002-11-14
; PRIOR APPLICATION NUMBER: US/09/321,589
; PRIOR FILING DATE: 1999-05-28
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1
; LENGTH: 735
; TYPE: PR1
; ORGANISM: Adeno-associated virus
US-10-293-478-1

Query Match 100.0%; Score 3994; DB 4; Length 735;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAADGYLPDWLEDTLSEGIQWVKLKPQPPPKPAERHKDSDRLVLPYKYLGPFGNGLD 60
DB 1 MAADGYLPDWLEDTLSEGIQWVKLKPQPPPKPAERHKDSDRLVLPYKYLGPFGNGLD 60
QY 61 KGEVNEADAAALSHDKAYDRQLDSGDNPLYKYNHADAERQERLKEDTSFGNIGRAVFQ 120
DB 61 KGEVNEADAAALSHDKAYDRQLDSGDNPLYKYNHADAERQERLKEDTSFGNIGRAVFQ 120
QY 121 AKKRVLEPLGLVEPVTAPGKKRPVSHSVPEPSSSGTGKAGQQAQKRLNFGQTGDAD 180
DB 121 AKKRVLEPLGLVEPVTAPGKKRPVSHSVPEPSSSGTGKAGQQAQKRLNFGQTGDAD 180
QY 181 SVDPQPLQGPAAAPSLGTNTMATGSGAPMADNNEGADVGNSSGNWHCDSTWMDRVI 240
DB 181 SVDPQPLQGPAAAPSLGTNTMATGSGAPMADNNEGADVGNSSGNWHCDSTWMDRVI 240
QY 241 TTSTRTRWALPTYNHLYKQISSQSGASNDNHYFGYSTPMGYFDNFRHCHFSRDPWRQRI 300
DB 241 TTSTRTRWALPTYNHLYKQISSQSGASNDNHYFGYSTPMGYFDNFRHCHFSRDPWRQRI 300
QY 301 NNNWGFRPKRLNFKLFIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQILPYVLGSAHQG 360

DB 301 NNNWGFRPKRLNFKLFIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQILPYVLGSAHQG 360
QY 361 CLPFPADVFMVPOYGYLTLLNGQAVGRSSFYCLEYFPSPQMLRTGNNTFSYTFEDVPF 420
DB 361 CLPFPADVFMVPOYGYLTLLNGQAVGRSSFYCLEYFPSPQMLRTGNNTFSYTFEDVPF 420
QY 421 HSSVAHSOSLRLANPLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
DB 421 HSSVAHSOSLRLANPLIDQYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
QY 481 PCYRQORVSKTSADNNSEYSGWTGATKYHLNGRDSLVPNPGPAMASHKDDKEKFFPQSGVL 540
DB 481 PCYRQORVSKTSADNNSEYSGWTGATKYHLNGRDSLVPNPGPAMASHKDDKEKFFPQSGVL 540
QY 541 IFGKGSEKTNVDIEKWMITDEBEIRTNPNVATEQYGSVSTNLQRNQAATAADVNTQGV 600
DB 541 IFGKGSEKTNVDIEKWMITDEBEIRTNPNVATEQYGSVSTNLQRNQAATAADVNTQGV 600
QY 601 LPGVMWQDRDVLQGPWAKIPHTDGHFHPSPLMGGFGLKHPHPQILIKNTVPVNPSTT 660
DB 601 LPGVMWQDRDVLQGPWAKIPHTDGHFHPSPLMGGFGLKHPHPQILIKNTVPVNPSTT 660
QY 661 FSAAKFASFTQYSTQGVSVSEIWELOKENSKEWNPFIQYTSNYSKSVNVDFTVDTNGVY 720
DB 661 FSAAKFASFTQYSTQGVSVSEIWELOKENSKEWNPFIQYTSNYSKSVNVDFTVDTNGVY 720
QY 721 SEPRPIGTRYLTRNL 735
DB 721 SEPRPIGTRYLTRNL 735

RESULT 3

US-09-807-802A-3
; Sequence 3, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNVEN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 3
; LENGTH: 736
; TYPE: PR1
; ORGANISM: AAV-1
US-09-807-802A-3

Query Match 85.2%; Score 3402.5; DB 4; Length 736;
Best Local Similarity 83.3%; Pred. No. 2-2e-276;
Matches 613; Conservative 51; Mismatches 71; Indels 1; Gaps 1;
QY 1 MAADGYLPDWLEDTLSEGIQWVKLKPQPPPKPAERHKDSDRLVLPYKYLGPFGNGLD 60
DB 1 MAADGYLPDWLEDTLSEGIQWVKLKPQPPPKPAERHKDSDRLVLPYKYLGPFGNGLD 60
QY 61 KGEVNEADAAALSHDKAYDRQLDSGDNPLYKYNHADAERQERLKEDTSFGNIGRAVFQ 120
DB 61 KGEVNEADAAALSHDKAYDRQLDSGDNPLYKYNHADAERQERLKEDTSFGNIGRAVFQ 120
QY 121 AKKRVLEPLGLVEPVTAPGKKRPVSHSVPEPSSSGTGKAGQQAQKRLNFGQTGDAD 180
DB 121 AKKRVLEPLGLVEPVTAPGKKRPVSHSVPEPSSSGTGKAGQQAQKRLNFGQTGDAD 180
QY 181 SVDPQPLQGPAAAPSLGTNTMATGSGAPMADNNEGADVGNSSGNWHCDSTWMDRVI 240

Matches 496; Conservative 43; Mismatches 59; Indels 1; Gaps 1;
QY 138 TAPGKQPVHSPVEPDPSSSGTGKAGQOPARKLNFGQGDADSVDPDPOPLGQPPAASG 197
Db 1 TAPGKQPVHSPVEPDPSSSGTGKAGQOPARKLNFGQGDADSVDPDPOPLGQPPAASG 60
QY 198 LGTNTMATGSGAPMADNNEGADGVGNSSGNHWCDSWTMGDRVITSTRTWALPTYNHLY 257
Db 61 VGPTTASGGGAPMADNNEGADGVGNSSGNHWCDSWTMGDRVITSTRTWALPTYNHLY 120
QY 258 KQISSOS-GASNDNHFGYSTPMGYDFNRFCHFSPRDWQRLINNNWGRPRKLNFKLP 316
Db 121 KQISSASTGASNDNHFGYSTPMGYDFNRFCHFSPRDWQRLINNNWGRPRKLNFKLP 180
QY 317 NIQKVEVTQNDGTTTIANNLTSVQVFTDSEYQLPVLSAHQGLCPFPFPADVPMVPOYG 376
Db 181 NIQKVEVTNDGTTTIANNLTSVQVFTDSEYQLPVLSAHQGLCPFPFPADVPMVPOYG 240
QY 377 YLTNNGSOAVGRSSFCLEYPPSQMLRTGNNFTSYTFEDVPFHSSYAHSQSLDRLMNP 436
Db 241 YLTNNGSOAVGRSSFCLEYPPSQMLRTGNNFTSYTFEDVPFHSSYAHSQSLDRLMNP 300
QY 437 LIDQLYLNRNTSGTTSQRLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNN 496
Db 301 LIDQLYLNRNTSGTTSQRLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNN 360
QY 497 NSEYSGTATKYHLNDRSLVNPMPAMASHKDDREKFFPQSGVLIIFGKQSEKTNVDIEK 556
Db 361 NSNFTWQASYNLNGRESIINPGTAMASHKDDREKFFPQSGVLIIFGKQSEKTNVDIEK 420
QY 557 VMTDEEIRTNVATEQYSGVSTNLQRGNRQAATADVNTQGLVPGMWQDRDVLQGP 616
Db 421 VMTDEEIRTNVATEQYSGVSTNLQRGNRQAATADVNTQGLVPGMWQDRDVLQGP 480
QY 617 IWAKIPHTDGHFHPSPLMGFGFLKHPPPPQILIKNTVPANPSTTFSAAKFAFSTQYSTG 676
Db 481 IWAKIPHTDGHFHPSPLMGFGFLKHPPPPQILIKNTVPANPSTTFSAAKFAFSTQYSTG 540
QY 677 QVSVEIEWELQKSKRWNPETQYTSNYSKNSVNVDTVDNTGVSSEPRIGTRYLTRL 735
Db 541 QVSVEIEWELQKSKRWNPETQYTSNYSKNSVNVDTVDNTGVSSEPRIGTRYLTRL 599
RESULT 6
US-09-807-802A-17
; Sequence 17, Application US/09807802A
; Patent No. 6759237
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; TITLE OF INVENTION: Xiao, Weidong
; TITLE OF INVENTION: Adeno-Associated Virus Serotype I Nucleic Acid Sequences,
; TITLE OF INVENTION: Vectors and Host Cells Containing Same
; FILE REFERENCE: GNVN.031USA
; CURRENT APPLICATION NUMBER: US/09/807,802A
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: US 60/107,114
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: PCT/US99/25694
; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 534
; TYPE: PRT
; ORGANISM: AAV-1
US-09-807-802A-17

Query Match
Best Local Similarity 62.1%; Score 2481.5; DB 4; Length 534;
Matches 445; Conservative 37; Mismatches 51; Indels 1; Gaps 1;
QY 203 MATGSGAPMADNNEGADGVGNSSGNHWCDSWTMGDRVITSTRTWALPTYNHLYKQISS 262

Db 1 MASGGGAPMADNNEGADGVGNASGNHWCDSWTMGDRVITSTRTWALPTYNHLYKQISS 60
QY 263 QS-GASNDNHFGYSTPMGYDFNRFCHFSPRDWQRLINNNWGRPRKLNFKLNIOVK 321
Db 61 ASGASNDNHFGYSTPMGYDFNRFCHFSPRDWQRLINNNWGRPRKLNFKLNIOVK 120
QY 322 EYTONDGTITIANNLTSVQVFTDSEYQLPVLSAHQGLCPFPFPADVPMVPOYGILTN 381
Db 121 EYTONDGTITIANNLTSVQVFTDSEYQLPVLSAHQGLCPFPFPADVPMVPOYGILTN 180
QY 382 NGSOAVGRSSFCLEYPPSQMLRTGNNFTSYTFEDVPFHSSYAHSQSLDRLMNPIDQY 441
Db 181 NGSOAVGRSSFCLEYPPSQMLRTGNNFTSYTFEDVPFHSSYAHSQSLDRLMNPIDQY 240
QY 442 LYLSRNTSGTTSQRLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNSEYS 501
Db 241 LYLSRNTSGTTSQRLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNSEYS 300
QY 502 WTGATKYHLNDRSLVNPMPAMASHKDDREKFFPQSGVLIIFGKQSEKTNVDIEKVMITD 561
Db 301 WTGATKYHLNDRSLVNPMPAMASHKDDREKFFPQSGVLIIFGKQSEKTNVDIEKVMITD 360
QY 562 EBEIRTNVATEQYSGVSTNLQRGNRQAATADVNTQGLVPGMWQDRDVLQGPIMAKI 621
Db 361 EBEIRTNVATEQYSGVSTNLQRGNRQAATADVNTQGLVPGMWQDRDVLQGPIMAKI 420
QY 622 PHTDGHFHPSPLMGFGFLKHPPPPQILIKNTVPANPSTTFSAAKFAFSTQYSTQVSVE 681
Db 421 PHTDGHFHPSPLMGFGFLKHPPPPQILIKNTVPANPSTTFSAAKFAFSTQYSTQVSVE 480
QY 682 IEWELQKSKRWNPETQYTSNYSKNSVNVDTVDNTGVSSEPRIGTRYLTRL 735
Db 481 IEWELQKSKRWNPETQYTSNYSKNSVNVDTVDNTGVSSEPRIGTRYLTRL 534
RESULT 7
US-09-532-594B-4
; Sequence 4, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.025202
; CURRENT APPLICATION NUMBER: US/09/532,594B
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 capsid protein VPI
US-09-532-594B-4

Query Match
Best Local Similarity 60.0%; Score 2397; DB 4; Length 734;
Matches 457; Conservative 90; Mismatches 167; Indels 36; Gaps 13;
QY 4 DGYLPDWEEDTISEGIROWWKLKPGPPPPKPAERHKDSDGLVLPQYKYPGPNGLDKGE 63
Db 3 DGYLPDWEEDTISEGIROWWKLKPGPPPPKPAERHKDSDGLVLPQYKYPGPNGLDKGE 62
QY 64 PVNEADAALHDKAYDRQLDSGDNPNLYKNHDAEFQERLKEDTSGGNLGRAVFOAKK 123
Db 63 PVNEADAALHDKAYDRQLDSGDNPNLYKNHDAEFQERLKEDTSGGNLGRAVFOAKK 122

```
QY 124 RVLEPLGLVEBPVKAPGKRVPVEHSPVDPDSSSTGKAGQOPARKRLNF-GQTGDADSV 182
DB 123 RVLEPLGLVEQAGETAPGKRPLIESPQPDSSSTGIGKKGQOPAKKLVFEDETGA GD-- 180
QY 183 PDPQLGQPPAAPSLG--GTTNTATGSGAPMADNNEGADGVGNSSGNWHCHDSTWMDRVI 240
DB 181 -----GPEGSTSGMNSDDSEKRAAAGAAVEGGQADGVGNAGSDWHCHDSTWSEGHVT 234
QY 241 TTSTRTWALPTYNHLYKOISSQAGASNDNHFGYSTPMGYDFDNRFHCHFSRDMQRLI 300
DB 235 TTSTRTWALPTYNHLYKRLGE---SLQNTYNGFSTPMGYDFDNRFHCHFSRDMQRLI 291
QY 301 NNNWGRPKRLNFKLNIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQG 360
DB 292 NNNWGRPKRPMKRVKFNIIQVKEVTTSNGETTVANNLTSTVQIFADSSYELPYVMDAGQSG 351
QY 361 CLPPPPADVFMVPOQGY---LTLNNGSAQVGRSSFYCLEYFPFSQMLRTGNNTFFSTYRED 417
DB 352 SLPPFPNDVFMVPOQGYCGLVNTGNTSQOQTDNRNPFYCLEYFPFSQMLRTGNNTFFSTYFEK 411
QY 418 VPFHSSYAHQSGLDRMLNPLIDQYLYLRTNTPSGTTTQ---SRLQFSQAGASDIRDQS 474
DB 412 VPFHSMYAHQSGLDRMLNPLIDQYLGWLGQSTTT--GTTLNAGTATTNFTKLRTNFSNFK 469
QY 475 RNWLFGPCVROQRVSKTSADNNSEYSWTGA---TKYH----LNGRDSLVPNPGPAMASHK 527
DB 470 KMWLPGPSIKQOGFSKTA--NQYKIPATGSDSLIKYETHSTLDGRWSALTGPPMATAG 527
QY 528 DDEEKFQSGVLIIG--KQSEKTNVDIEKVMITDEBEIRTNPNVATEQYGSVSTNLR 585
DB 528 PADSKF--SNSQLIFAGPKQNGTATVP--GTLIFTSEELAATNATDTDMWGNLPGDQGS 584
QY 586 GNRQAATAADVNTQGVLPQVMWQDRDVIYQGPWAKIPHTDGHFSPMLGGFGLKHPFPQ 645
DB 585 NSNLPTVDRLTALGAVPGVMWQNRDIYQGPWAKIPHTDGHFSPMLGGFGLKHPFPQ 644
QY 646 ILIKNTVPANPSTTFSSAAKFASTQYSTQGVSVIEWELQKENSKRWNPIQVTSNN 705
DB 645 IPKNTVPANPATTFSSTPVNSFTQYSTQGVSVQIDWEIQKERSKRWNPEVQTSNYG 704
QY 706 KSVNVDFVTDTNGVYSEPRPICTRYLTRNL 735
DB 705 QNSLWAPDAAGKYTEPRAIGTRYLTRNL 734
```

RESULT 8

```
US-09-532-594B-16
; Sequence 16, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.025202
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 598
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; OTHER INFORMATION: synthetic construct
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 capsid protein VP2
US-09-532-594B-16
```

Query Match 45.0%; Score 1797; DB 4; Length 598;
Best Local Similarity 56.7%; Pred. No. 6.9e-142;

```
Matches 349; Conservative 76; Mismatches 155; Indels 36; Gaps 13;
QY 138 TAPGKRVPVEHSPVDPDSSSTGKAGQOPARKRLNF-GQTGDADSVDPQPLGQPPAAPS 196
DB 1 TAPGKRPLIESPQPDSSSTGIGKKGQOPAKKLVFEDETGA GD-----GPEGSTS 52
QY 197 GL--GTTNTATGSGAPMADNNEGADGVGNSSGNWHCHDSTWMDRVIITSTRTWALPTYN 254
DB 53 GMSDDSEKRAAAGAAVEGGQADGVGNAGSDWHCHDSTWSEGHVTITSTRTWALPTYN 112
QY 255 HLYKOISSQAGASNDNHFGYSTPMGYDFDNRFHCHFSRDMQRLINNNWGRPKRLNFK 314
DB 113 HLYKRLGE---SLQNTYNGFSTPMGYDFDNRFHCHFSRDMQRLINNNWGRPKRMRVK 169
QY 315 LFNIIQVKEVTQNDGTTTIANNLSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPO 374
DB 170 LFNIIQVKEVTTSNGETTVANNLTSTVQIFADSSYELPYVMDAGQSGSLPPFPNDVFMVPO 229
QY 375 YGY---LTLNNGSAQVGRSSFYCLEYFPFSQMLRTGNNTFFSTYFEDVFPFHSSYAHQSGLD 431
DB 230 YGYCGLVNTGNTSQOQTDNRNPFYCLEYFPFSQMLRTGNNTFFSTYFEDVFPFHSSYAHQSGLD 289
QY 432 RLNNPLIDQYLYLRTNTPSGTTTQ---SRLQFSQAGASDIRDQSRNWLPGPCYRQORV 488
DB 230 RLNNPLIDQYLGWLGQSTTT--GTTLNAGTATTNFTKLRTNFSNFKMWLPGPSIKQOGF 347
QY 489 SKTSADNNSEYSWTGA---TKYH----LNGRDSLVPNPGPAMASHKDDEEKFQSGVLI 541
DB 348 SKTA--NQYKIPATGSDSLIKYETHSTLDGRWSALTGPPMATAGPADSKF--SNSQLI 403
QY 542 FG--KQSEKTNVDIEKVMITDEBEIRTNPNVATEQYGSVSTNLRQNRQAATAADVNTQ 599
DB 404 FAGPKQNGTATVP--GTLIFTSEELAATNATDTDMWGNLPGDQGSNLTPTVDRLTALG 462
QY 600 VLPQVMWQDRDVIYQGPWAKIPHTDGHFSPMLGGFGLKHPFPQILIKNTVPANPST 659
DB 463 AVPGVMWQNRDIYQGPWAKIPHTDGHFSPMLGGFGLKHPFPQIFIKNTVPANPAT 522
QY 660 TFSAAKFASTQYSTQGVSVIEWELQKENSKRWNPIQVTSNNKSVNVDFVTDTNGV 719
DB 523 TFSSTPVSFTQYSTQGVSVQIDWEIQKERSKRWNPEVQTSNYGQNSLWAPDAAGK 582
QY 720 YSEPRPICTRYLTRNL 735
DB 583 YTEPRAIGTRYLTRNL 598
```

RESULT 9

```
US-09-532-594B-18
; Sequence 18, Application US/09532594B
; Patent No. 6468524
; GENERAL INFORMATION:
; APPLICANT: Chorini, John A.
; APPLICANT: Kotin, Robert M.
; APPLICANT: Safer, Brian
; APPLICANT: Davidson, Beverly
; TITLE OF INVENTION: AAV4 VECTOR AND USES THEREOF
; FILE REFERENCE: 14014.025202
; CURRENT APPLICATION NUMBER: US/09/532,594B
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6468524e =
; OTHER INFORMATION: synthetic construct
; NAME/KEY: misc feature
; OTHER INFORMATION: AAV4 capsid protein VP3
US-09-532-594B-18
```


COMPUTER: AT&T - IBM COMPATIBLE
OPERATING SYSTEM: MS-DOS Version 6.2
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/856,841
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/214,658
FILING DATE: 16-MARCH-1994
APPLICATION NUMBER: US 07/917,096
FILING DATE: 4-AUGUST-1992
APPLICATION NUMBER: PCT/DE91/00106
FILING DATE: 8-FEBRUARY-1991
APPLICATION NUMBER: DE40038262
FILING DATE: 8-FEBRUARY-1990
ATTORNEY/AGENT INFORMATION:
NAME: ROBINSON, WILLIAM R.
REGISTRATION NUMBER: 27,224
REFERENCE/DOCKET NUMBER: LKR-9222-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 543
TYPE: AMINO ACID
TOPOLOGY: LINEAR
MOLECULE TYPE: PEPTIDE
DESCRIPTION: PEPTIDE
HYPOTHETICAL: N/A
ANTI-SENSE: N/A
FRAGMENT TYPE: INTERNAL
ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
IMMEDIATE SOURCE: INFECTION (ERYTHEMA INFECTION)
POSITION IN GENOME: N/A
FEATURE:
NAME/KEY: N/A
LOCATION: N/A
IDENTIFICATION METHOD: amino acid analysis and
IDENTIFICATION METHOD: mass spectrometry
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: COSSART, Y.E.
AUTHORS: FIELD, A.M.
AUTHORS: CANT, B.
AUTHORS: WIDDOWS, D.
TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
JOURNAL: LANCET
VOLUME: I
ISSUE:
PAGES: 72 - 73
DATE: 1975
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO: 22:
US-08-856-841-22

Query Match 12.2%; Score 486; DB 3; Length 543;
Best Local Similarity 27.4%; Pred. No. 5,1e-32;
Matches 148; Conservative 80; Mismatches 221; Indels 92; Gaps 17;
Qy 202 TWATGS----GAPMADNNEGADGVNSS--GNWHCDSTWMDGRVITTTTRTWALPTYNH 255
Db 2 TWITNSLIMTSVNSAEASTGAGGGGSGNSVKMSWSEGAFTSANSVTCFTSRQFLIPYDPEH 61
Qy 256 LYKQISSQSGASND-----NHVFGYSTPMGYDFNRFCHFSRDMQRLNNW 304
Db 62 HYKVFSPAASCHNAGSKEAKVCTTSPMGYSTPWRYLDFNALNLFPSLRFQHLIENYG 121

RESULT 12

US-08-856-841-16
; Sequence 16, Application US/08856841
; Patent No. 6274307
; GENERAL INFORMATION:
; APPLICANT: ERWIN SOUTSCHEK
; APPLICANT: MANFRED MOTZ
; TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
; TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
; STREET: 99 PARK AVENUE
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" FLOPPY DISC
; COMPUTER: AT&T - IBM COMPATIBLE
; OPERATING SYSTEM: MS-DOS Version 6.2
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856,841
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/214,658
; FILING DATE: 16-MARCH-1994
; APPLICATION NUMBER: US 07/917,096
; FILING DATE: 4-AUGUST-1992
; APPLICATION NUMBER: PCT/DE91/00106
; FILING DATE: 8-FEBRUARY-1991
; APPLICATION NUMBER: DE40038262
; FILING DATE: 8-FEBRUARY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINSON, WILLIAM R.
; REGISTRATION NUMBER: 27,224
; REFERENCE/DOCKET NUMBER: LKR-9222-A
; TELECOMMUNICATION INFORMATION:

Qy 305 GFRPKRLNFKLFNIQKVEVT-QNDGTTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLP 363
Db 122 SIAPDALTVTISEIAVKDVTDKTGGGVQVTDSTTGRCLCMLVDHEVKYPYVLGGQDITLAP 181
Qy 364 PFPADVFMVPOYGYLTLAN-LSQAVG-----RSSFYCLEYFSPQMLRTGNFTFSY 413
Db 182 ELPIWVYFPQYAYLTGVDVNTQGISGDSKKLASEESAFYVLEHSSFQLLGTGTASMSY 241
Qy 414 TFDVPFHSSVAHSQSLORLNNPLIDQVLYLSTRNTSPGTTQSRLOFSOAGASDIRDQ 473
Db 242 KFPVPPELENLCCSQHFYEMYNPL-----YCSRLGVDP--TLGGDPKFRSLTHEDHAIQ 293
Qy 474 SRNMLPGPCYRQVRKTSADNNSEYSWTG-ATKYHLNGRDSLVPNPGP-AWASHKDBEE 531
Db 294 PQNFMGGLVNSVSTKEGDSNTGAKALTGLSTGTSQNTRI-SL-RPGVSPQYHHWDTD 352
Qy 532 KFFPQSGVLIETG-----KQSEKTNVDIEKVMITDEEIRITNP-VATEQVGS 578
Db 353 KYVTGINAISHGQTTYGNAEDKEYOQGVGRFPNEKEBQLKQLQGLNMHTYFPNKGTOQY-- 410
Qy 579 VSTNLQRGNRQAATADVNTQGVLPQGMVWQDRDQVYLGQPIWAKIPHTDGHFHP-PLMGGF 637
Db 411 -----TDQIERPLMVGSVWNRALHYESQLWSKIPNLDSDSKTQFPAALGOW 456
Qy 638 GLKHPPQILIKNTVPANPSTTFSAAKFAFITQYSTQGVSVIEWEL-QKENSKRWNP 696
Db 457 GLHQPPQI-----FLQYAVGIMTVTMTFKLGPRAKATGRWNP 494
Qy 697 E 697
Db 495 Q 495


```

; TELEPHONE: (212) 697-3355
; TELEFAX: (212) 557-5635
; TELEX: NONE
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 501
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PEPTIDE
; DESCRIPTION: PEPTIDE
; HYPOTHETICAL: N/A
; ANTI-SENSE: N/A
; FRAGMENT TYPE: INTERNAL
; ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
; IMMEDIATE SOURCE: INFECTION (ERYTHEMA INFECTIONIUM)
; POSITION IN GENOME: GENETICALLY ENGINEERED PEPTIDE
; FEATURE:
; NAME/KEY: N/A
; LOCATION: N/A
; IDENTIFICATION METHOD: amino acid analysis and
; IDENTIFICATION METHOD: mass spectrometry
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: COSSART, Y.E.
; AUTHORS: FIELD, A.M.
; AUTHORS: CANT, B.
; TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
; JOURNAL: LANCET
; VOLUME: 1
; ISSUE:
; PAGES: 72 - 73
; DATE: 1975
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO: 18:
; US-08-856-841-18

Query Match 10.8%; Score 431; DB 3; Length 501;
Best Local Similarity 27.5%; Pred. No. 1.8e-27;
Matches 116; Conservative 50; Mismatches 180; Indels 76; Gaps 7;

Qy 38 HKDSSGLVPGYKYLGPFGNGLDKEPVNEADAAALEHDKAYDRQLDSDGNPYLKYNHAD 97
Db 125 HKPGQVSQVLPCTNYVGGNELQAGPPQSAVDSAAIHDFRYSQAKLGINPYTHWTVD 184
Qy 98 AEFQRLKEDTSFGNLRGAVPQAKRVLLEPLGLVEEPVKTAGKRPVHSPVPEPDSS 157
Db 185 BELKNIKNETGFQAVVKDYF-----TLKGAAAPVAHP----- 218
Qy 158 GTGKAGQAPARKLNFQGTGDADSVDPDQPLGQPPAAPSGLTNTMTATGSGAPMADNNEG 217
Db 219 -----QGSUPEVPAYNASEKYPSTSVNSAEASTGA-----G 250
Qy 218 ADGVGNSSGNHCDSTWMDRVITTSRTWALPTYNNHLYKQISSQSGASND----- 269
Db 251 CGGSNSVSKMSWSEGATFSANSTCTFSRQFLIPYDEHYKVFSPAASCHNASKEAKV 310
Qy 270 ---NHFGYSTPWGYDFNRHFCHFSPRDWQLINNNWGFPRKRLNFKLFNIQKVEVT-Q 325
Db 311 CTISPIMGYSTPWRYLDFNALNLFSPLEFQHLIENYGSIAPDALTVTISETAVKDVTDK 370
Qy 326 NDTGTTIANNLTSTVOVTFDSEYQLPYVLGSAHQCLPPFPADVPWVQYGLTLNN-GS 384
Db 371 TGGGVQVTDSTTGRICMLVDHEYKYPYVLGQGGQDTLAPELPIWVYFPQYALTYGDVNT 430
Qy 385 QAVG-----RSSFCLEYFPQMLRTGNFTFSYTFEDVPFPHSSVAHSQSQSLRLMN 435
Db 431 QGISGSKKLASESAFVLEHSSFQLLGTGTGASMSYKFPVPENLEGCSQHPEMYN 490
Qy 436 PL 437

; Db 491 PL 492
;
; RESULT 14
; US-08-856-841-19
; Sequence 19, Application US/08856841
; Patent No. 6274307
; GENERAL INFORMATION:
; APPLICANT: ERWIN SOUTSCHEK
; APPLICANT: MANFRED MOTZ
; TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
; TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
; STREET: 99 PARK AVENUE
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" FLOPPY DISC
; COMPUTER: AT&T - IBM COMPATIBLE
; OPERATING SYSTEM: MS-DOS Version 6.2
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856,841
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/214,658
; FILING DATE: 16-MARCH-1994
; APPLICATION NUMBER: US 07/917,096
; FILING DATE: 4-AUGUST-1992
; APPLICATION NUMBER: PCT/DE91/00106
; FILING DATE: 8-FEBRUARY-1991
; APPLICATION NUMBER: DE40038262
; FILING DATE: 8-FEBRUARY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINSON, WILLIAM R.
; REGISTRATION NUMBER: 27,224
; REFERENCE/DOCKET NUMBER: LKR-9222-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-3355
; TELEFAX: (212) 557-5635
; TELEX: NONE
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 486
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PEPTIDE
; DESCRIPTION: PEPTIDE
; HYPOTHETICAL: N/A
; ANTI-SENSE: N/A
; FRAGMENT TYPE: INTERNAL
; ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
; IMMEDIATE SOURCE: INFECTION (ERYTHEMA INFECTIONIUM)
; POSITION IN GENOME: GENETICALLY ENGINEERED PEPTIDE
; FEATURE:
; NAME/KEY: N/A
; LOCATION: N/A
; IDENTIFICATION METHOD: amino acid analysis and
; IDENTIFICATION METHOD: mass spectrometry
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: COSSART, Y.E.
; AUTHORS: FIELD, A.M.
; AUTHORS: CANT, B.
; TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
```

RESULT 15

Db 184 NTRISL-REGPVSQPYHHWDTDKYVTGINAISHGQTTYGNAEDKEYQQGVGRFPNEKEOL 242
QY 558 MITDEEEIRTTNP-VATEOYGSVSTNLQRGNQAAATADVNTQGVLPGMVWQDRDVVLOGP 616
Db 243 KOLOGLNWHITYFPNKCTQY-----TDQIERPLMVGSVWNRRAHYESQ 286
QY 617 IWAKIPHTDGHFHS-PLMGFGLKHPPPOILIKNTPVPANPSTTFSAAKFAFIT--QY 673
Db 287 LWSKIPNLDDSFKQFAALGGWGLHQPPPOIFLKILP-----ESGPIGGIKSMGITTLOVY 342
QY 674 STGOVSVEIEWEL-OKENSKRWNP 697
Db 343 AVGIMTVTTFKLGPRKATGRWNPO 367

Search completed: May 19, 2005, 13:34:37
Job time : 44 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2005, 13:15:05 ; Search time 165 Seconds
(without alignments)
1722.842 Million cell updates/sec

Title: US-10-038-972A-13

Perfect score: 3994

Sequence: 1 MAADGYLPDWLEDTLSEGR.....TNGVSEPRPIGTRYLTRNL 735

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3994	100.0	735	4	AAG65792 Adeno-ss
2	3994	100.0	735	4	AAB59844 AAV2 caps
3	3994	100.0	735	4	AAM51508 Adeno-ss
4	3994	100.0	735	5	AAY98974 Adeno-ss
5	3994	100.0	735	6	ABR80231 AAV2 vpl
6	3994	100.0	735	6	ABR82436 Adeno-ss
7	3994	100.0	735	7	ABR62761 Adeno-ss
8	3994	100.0	735	7	ADP76571 Adeno-ss
9	3967.5	99.3	734	4	AAB50326 Adeno-ss
10	3623	90.7	735	7	ADP76568 Adeno-ss
11	3622	90.7	735	7	ADP76570 Adeno-ss
12	3617	90.6	735	7	ADP76569 Adeno-ss
13	3615	90.5	735	7	ADP76567 Adeno-ss
14	3557.5	89.1	736	4	AAB59846 AAV3B cap
15	3532.5	88.4	736	4	AAB59845 AAV3A cap
16	3532.5	88.4	736	6	ABR80233 AAV3 vpl
17	3532.5	88.4	736	7	ABR62763 Adeno-ss
18	3532.5	88.4	736	7	ADP76572 Adeno-ss
19	3421	85.7	731	7	ADP76589 Adeno-ss
20	3416.5	85.5	738	7	ADP76592 Adeno-ss
21	3409.5	85.4	736	4	AAB59847 AAV6 caps
22	3409.5	85.4	736	7	ADP76566 Adeno-ss
23	3405.5	85.3	738	7	ADP76595 Adeno-ss
24	3405.5	85.3	738	7	ADP76594 Adeno-ss
25	3404.5	85.2	738	7	ADP76580 Adeno-ss

26	3404.5	85.2	738	7	ADP76586	Ade76586 Adeno-ss
27	3404.5	85.2	738	7	ADP76582	Ade76582 Adeno-ss
28	3402.5	85.2	736	3	AAV71167	AAV71167 Adeno-ss
29	3402.5	85.2	736	6	ABR80232	ABR80232 AAV1 vpl
30	3402.5	85.2	736	7	ABR62762	ABR62762 Adeno-ss
31	3402.5	85.2	736	7	ADP76565	Ade76565 Adeno-ss
32	3401	85.2	731	7	ADP76590	Ade76590 Adeno-ss
33	3398.5	85.1	738	7	ADP76593	Ade76593 Adeno-ss
34	3397.5	85.1	738	7	ADP76581	Ade76581 Adeno-ss
35	3396	85.0	733	7	ADP76591	Ade76591 Adeno-ss
36	3392.5	84.9	736	7	ADP76598	Ade76598 Adeno-ss
37	3392.5	84.9	738	7	ADP76583	Ade76583 Adeno-ss
38	3391.5	84.9	736	6	ABR80229	ABR80229 AAV9 cap
39	3391.5	84.9	736	7	ABR62764	ABR62764 Adeno-ss
40	3391.5	84.9	736	7	ADP76601	Ade76601 Adeno-ss
41	3391.5	84.9	738	7	ADP76585	Ade76585 Adeno-ss
42	3386.5	84.8	738	7	ADP76584	Ade76584 Adeno-ss
43	3383.5	84.7	736	7	ADP76597	Ade76597 Adeno-ss
44	3381	84.7	733	7	ADP76588	Ade76588 Adeno-ss
45	3379.5	84.6	736	7	ADP76599	Ade76599 Adeno-ss

ALIGNMENTS

RESULT 1

AAG65792

ID AAG65792 standard; protein; 735 AA.

XX AC AAG65792;

XX DT 11-FEB-2002 (first entry)

XX DB Adeno-associated virus 2 (AAV-2) major coat protein VPI.

XX KW Recombinant viral vector; RVV; capsid; parvovirus; transgene; cytostatic; inverted terminal repeat; neurotropic; antitumor; gene therapy; adeno-associated virus; AAV; major coat protein; AAV-2; VPI.

XX OS Adeno-associated virus 2.

XX PN WO200168888-A2.

XX PD 20-SEP-2001.

XX PF 13-MAR-2001; 2001WO-US007927.

XX PR 14-MAR-2000; 2000US-0189110P.

XX PA (NEUR-) NEUROLOGIX INC.

XX PI Xiao W, During MJ;

XX DR WPI; 2001-596912/67.

XX N-PSDB; AAI66974.

XX PT Recombinant viral vector useful in improving gene therapy in a subject, and for increasing efficiency of entry into a cell, comprises a chimeric capsid having one non-native amino acid sequence and a desired transgene.

XX PS Disclosure; Page 50; 53pp; English.

XX CC The invention provides a recombinant viral vector (RVV) comprising a chimeric capsid (I) having at least one non-native amino acid sequence, derived from a capsid protein domain of parvovirus (II), a virus (III), or their combination, and a transgene flanked 5' and 3' by inverted terminal repeat (ITR) sequences, derived from (II), (III), or their combination. The RVV is useful for improving gene therapy in a subject with a disorder, and for increasing the efficiency of entry into a cell, which involves producing (I) encapsulating a viral vector, and contacting a cell with RVV having (I) such that (I) binds to an attachment site on the cell surface and permits the vector to enter the cell efficiently. A

091304 896

pharmaceutical composition comprising RVV with (I) containing a transgene sequence associated with a disease or a disorder such that expression of the transgene would result in amelioration of the disease or disorder such as inherited neurological and metabolic diseases e.g. lysosomal storage disease, Lesch-Nyhan syndrome, amyloid polynuropathy, Alzheimer amyloid, Duchenne's muscular dystrophy, diseases of the blood e.g. sickle cell anemia, clotting disorders and thalassemias, cystic fibrosis, diabetes, diseases associated with hormone deficiencies, cystic fibrosis, and various types of neoplastic cells which include tumours especially central nervous system tumours, neoplasms, carcinomas, sarcomas, leukemias and lymphoma. The present sequence represents the adeno-associated virus 2 (AAV-2) major coat protein Vp1. AAV-2 sequences are used in the construction of a chimeric vector

Query Match 100.0%; Score 3994; DB 4; Length 735;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADGVLDPWLEDTLSEGIQWKKLPGPPPKPAERHKDDSRGLVLPYKYLGPFGNGLD 60
DB 1 MAADGVLDPWLEDTLSEGIQWKKLPGPPPKPAERHKDDSRGLVLPYKYLGPFGNGLD 60

QY 61 KGEVNEADAAALEHDKAYDRQLDSDGNPILKYNHADAFAERLKEDETSFGNLRGAVFQ 120
DB 61 KGEVNEADAAALEHDKAYDRQLDSDGNPILKYNHADAFAERLKEDETSFGNLRGAVFQ 120

QY 121 AKKRVLEPLGLVEEPVKTAPGKRPVHSVPEDSSGTGKAGQOPARKRLNFGQTGDAD 180
DB 121 AKKRVLEPLGLVEEPVKTAPGKRPVHSVPEDSSGTGKAGQOPARKRLNFGQTGDAD 180

QY 181 SVDPDQPLQPPAAPSGLTNTMATGSGAPMADNNEGADGVNSGNWHCDSTWMDRVI 240
DB 181 SVDPDQPLQPPAAPSGLTNTMATGSGAPMADNNEGADGVNSGNWHCDSTWMDRVI 240

QY 241 TTSTRTWALPTNNHLYKQISSQSGASNDNHFGYSTPWFYDFNRFHCHFSPRDWORLI 300
DB 241 TTSTRTWALPTNNHLYKQISSQSGASNDNHFGYSTPWFYDFNRFHCHFSPRDWORLI 300

QY 301 NNNWGFPRKRLNFKLFIQVKEVTQNDGTTIANNLTSTVQVFTDSEYQLPYLGSAGHQ 360
DB 301 NNNWGFPRKRLNFKLFIQVKEVTQNDGTTIANNLTSTVQVFTDSEYQLPYLGSAGHQ 360

QY 361 CLPPPADVFMVPPQGYLTLLNGSQAVGRSSFCYCLEYFPPSQMLRTGNNFTFSYTFEDVPF 420
DB 361 CLPPPADVFMVPPQGYLTLLNGSQAVGRSSFCYCLEYFPPSQMLRTGNNFTFSYTFEDVPF 420

QY 421 HSSVAHSQSLDRLMNPILDQYLYLSTRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
DB 421 HSSVAHSQSLDRLMNPILDQYLYLSTRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480

QY 481 PCYQQRVSKTSADNNNSEYSWTQATKYHLNGRSLVNPMPAMASHKDDSEKFFPQSGVL 540
DB 481 PCYQQRVSKTSADNNNSEYSWTQATKYHLNGRSLVNPMPAMASHKDDSEKFFPQSGVL 540

QY 541 IFGQSGSEKTNVDIEKWIITDEEIRTNVATEQYGSVSTNLQRGNRQAATADVNTQGV 600
DB 541 IFGQSGSEKTNVDIEKWIITDEEIRTNVATEQYGSVSTNLQRGNRQAATADVNTQGV 600

QY 601 LPMVMDQRDYLQGPVIAKIPHTDGHFHPSPLMGGFGLKHPPPQILLIKNTVPANPSIT 660
DB 601 LPMVMDQRDYLQGPVIAKIPHTDGHFHPSPLMGGFGLKHPPPQILLIKNTVPANPSIT 660

QY 661 FSAKFAFSTQYSTGVSEIWELOKENSQKWNPEIQTSYNKSVNVDFVDTNGVY 720
DB 661 FSAKFAFSTQYSTGVSEIWELOKENSQKWNPEIQTSYNKSVNVDFVDTNGVY 720

QY 721 SEPRPIGTRILTRNL 735
DB 721 SEPRPIGTRILTRNL 735

RESULT 2
AAB59844
ID AAB59844 standard; protein; 735 AA.
AC AAB59844;
XX
DT 28-MAR-2001 (first entry)
XX
DE AAV2 capsid protein Vp1.
XX
KW AAV2; gene therapy; AAV viral vector; cystic fibrosis; cancer; AIDS;
KW atherosclerosis; sickle cell anaemia; thalassemia;
KW blood clotting disorder; diabetes; capsid protein Vp1.
XX
OS Adeno associated virus.
XX
PN US6156303-A.
XX
PD 05-DEC-2000.
XX
PF 11-JUN-1997; 97US-00873168.
XX
PR 11-JUN-1997; 97US-00873168.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Russell DW, Rutledge EA;
XX
DR WPI; 2001-060164/07.
XX
PT Adeno-associated virus serotype 6 and viral vector derived from it for
PT gene therapy of cystic fibrosis, cancer, acquired immunodeficiency
PT syndrome, sickle cell anemia, thalassemia and diabetes.
XX
PS Claim 7; Fig 2; 50pp; English.
XX
CC The present invention relates to adeno-associated virus serotypes. The
CC present sequence is capsid protein Vp1 of one such serotype (AAV2). AAV2
CC can be used to construct AAV viral vectors for use in gene therapy for a
CC range of disorders: cystic fibrosis, cancer, AIDS, atherosclerosis,
CC sickle cell anaemia, thalassemia, blood clotting disorders and diabetes.
CC The AAV viral vectors have increased transduction efficiency of a
CC particular host cell as the AAV virion containing the AAV vector genome
CC can be modified to express a capsid protein of an AAV serotype that
CC transduces the selected host cell
XX
SQ Sequence 735 AA;

Query Match 100.0%; Score 3994; DB 4; Length 735;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADGVLDPWLEDTLSEGIQWKKLPGPPPKPAERHKDDSRGLVLPYKYLGPFGNGLD 60
DB 1 MAADGVLDPWLEDTLSEGIQWKKLPGPPPKPAERHKDDSRGLVLPYKYLGPFGNGLD 60

QY 61 KGEVNEADAAALEHDKAYDRQLDSDGNPILKYNHADAFAERLKEDETSFGNLRGAVFQ 120
DB 61 KGEVNEADAAALEHDKAYDRQLDSDGNPILKYNHADAFAERLKEDETSFGNLRGAVFQ 120

QY 121 AKKRVLEPLGLVEEPVKTAPGKRPVHSVPEDSSGTGKAGQOPARKRLNFGQTGDAD 180
DB 121 AKKRVLEPLGLVEEPVKTAPGKRPVHSVPEDSSGTGKAGQOPARKRLNFGQTGDAD 180

QY 181 SVDPDQPLQPPAAPSGLTNTMATGSGAPMADNNEGADGVNSGNWHCDSTWMDRVI 240
DB 181 SVDPDQPLQPPAAPSGLTNTMATGSGAPMADNNEGADGVNSGNWHCDSTWMDRVI 240

QY 241 TTSTRTWALPTNNHLYKQISSQSGASNDNHFGYSTPWFYDFNRFHCHFSPRDWORLI 300
DB 241 TTSTRTWALPTNNHLYKQISSQSGASNDNHFGYSTPWFYDFNRFHCHFSPRDWORLI 300

QY 301 NNNWGFPRKRLNFKLFIQVKEVTQNDGTTIANNLTSTVQVFTDSEYQLPYLGSAGHQ 360

PN WO200253703-A2.
 PD 11-JUL-2002.
 PP 04-JAN-2002; 2002WO-US000152.
 PR 05-JAN-2001; 2001US-0260124P.
 PA (CHIL-) CHILDRENS HOSPITAL INC.
 PI Bartlett JS;
 XX WPI; 2002-583608/62.
 DR N-P8DB; ABK89694.
 XX
 PT New adeno-associated virus vector comprises a biotinylated capsid or
 PT capsid protein with an amino acid insertion in the vp1 capsid, useful as
 PT a vaccine or for transferring a therapeutic peptide to a cancer cell.
 PS Claim 1; Page 49-51; 57pp; English.
 XX
 CC The invention relates to an adeno-associated virus (AAV) vector (I)
 CC comprising a biotinylated capsid or capsid protein (II) with an amino
 CC acid insertion following the capsid amino acid at position 139, 161, 588
 CC or 657 in the vp1 capsid. The AAV vector comprises a capsid protein
 CC containing one or more amino acid insertions that ablate the ability of
 CC the vector to bind heparin-sulphate proteoglycan and allow the vector to
 CC use a cellular receptor not used by wild type AAV. Modified (II) are
 CC useful as vaccines to elicit immune responses to amino acids, where the
 CC response can be protective and/or therapeutic. (I) may be used to
 CC transfer a therapeutic peptide to a cancer cell, particularly to an
 CC ovarian cancer cell. The present sequence represents the adeno-associated
 CC virus 2 (AAV2) vector, vp1 capsid protein used in to make modified AAV2
 CC vectors
 XX
 SQ Sequence 735 AA;
 Query Match 100.0%; Score 3994; DB 5; Length 735;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAADGYLPDWLEDTLSEGIROWWKLKPGPPPPKPAERHKDPSRLGLVLPVGYKLGPNGLD 60
 Db 1 MAADGYLPDWLEDTLSEGIROWWKLKPGPPPPKPAERHKDPSRLGLVLPVGYKLGPNGLD 60
 QY 61 KGEVNEADAALRHDKAYDQLSDGNDPNLYKYNHADAEPQERLKEDTSFGNLRGAVFQ 120
 Db 61 KGEVNEADAALRHDKAYDQLSDGNDPNLYKYNHADAEPQERLKEDTSFGNLRGAVFQ 120
 QY 121 AKKRVLEPLGLVEEPVKTAPGKRPVHSVPEDSSSGTGKAGQOPARKELNFGQTGDAD 180
 Db 121 AKKRVLEPLGLVEEPVKTAPGKRPVHSVPEDSSSGTGKAGQOPARKELNFGQTGDAD 180
 QY 181 SVDPDQPLGQPPAPSGLGTNTMTATGSGAPMADNEGAGVSGNWHCDSTWMDRVI 240
 Db 181 SVDPDQPLGQPPAPSGLGTNTMTATGSGAPMADNEGAGVSGNWHCDSTWMDRVI 240
 QY 241 TTSRTWALPTNNHLYKQISSQSGASNDNHYGYSTPMGYDFNPFCHPSRDWQRLI 300
 Db 241 TTSRTWALPTNNHLYKQISSQSGASNDNHYGYSTPMGYDFNPFCHPSRDWQRLI 300
 QY 301 NNNWGPRLNFKLNIQVKEVTQNDGTTIANNITSTVQVFTDSEYQLPYVLGSAHQG 360
 Db 301 NNNWGPRLNFKLNIQVKEVTQNDGTTIANNITSTVQVFTDSEYQLPYVLGSAHQG 360
 QY 361 CLPPFPADVPMVQYGYLTINNGSQVGRSSFYCLEYFPPSQMLRTGNFTSYTETDVPF 420
 Db 361 CLPPFPADVPMVQYGYLTINNGSQVGRSSFYCLEYFPPSQMLRTGNFTSYTETDVPF 420
 QY 421 HSSVAHSQSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOPSQAGASDIRDQSRNLWLP 480
 Db 421 HSSVAHSQSLDRMLNPLIDQYLYLSRTNTPSGTTTQSRLOPSQAGASDIRDQSRNLWLP 480

QY 481 PCYRQORVSKTSADNNSEYSWTGATKYHLNGRDSLVPMPAMASHKDDDEKRPQSGVL 540
 Db 481 PCYRQORVSKTSADNNSEYSWTGATKYHLNGRDSLVPMPAMASHKDDDEKRPQSGVL 540
 QY 541 IFGKGSGSEKTNVDIEKWMITDEEIRITTNVATQYGSVSTNLRGNRQAATADVNTQGV 600
 Db 541 IFGKGSGSEKTNVDIEKWMITDEEIRITTNVATQYGSVSTNLRGNRQAATADVNTQGV 600
 QY 601 LPGMVQODRDVYLOQPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
 Db 601 LPGMVQODRDVYLOQPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
 QY 661 FSAAKFASFTQYSTGVSEIWELOKENSKNWNEIQTYSNYSKSNVNDFTVDTNGVY 720
 Db 661 FSAAKFASFTQYSTGVSEIWELOKENSKNWNEIQTYSNYSKSNVNDFTVDTNGVY 720
 QY 721 SEPRPICTRYLTRNL 735
 Db 721 SEPRPICTRYLTRNL 735
 RESULT 5
 ABB80231
 ID ABB80231 standard; protein; 735 AA.
 XX ABB80231;
 AC ABB80231;
 XX 20-NOV-2003 (first entry)
 DT AAV2 vp1 protein.
 DE AAV2 vp1 protein.
 XX Adeno-associated virus; AAV; serotype 9; rep; cap; vp1; vp2; vp3;
 KW splice variant; transgene.
 XX Adeno associated virus serotype 2.
 OS WO2003052052-A2.
 PN 26-JUN-2003.
 PD 12-NOV-2002; 2002WO-US033631.
 XX 17-DEC-2001; 2001US-0341150P.
 PR 05-JUN-2002; 2002US-0386132P.
 XX (UYPE-) UNIV PENNSYLVANIA.
 PA Gao G, Wilson JM, Alvira M;
 PI WPI; 2003-523523/49.
 DR New isolated adeno-associated virus (AAV) comprising an AAV9 capsid,
 PT useful for preparing a medicament for delivering a transgene to a cell.
 XX Disclosure; Fig 2; 76pp; English.
 CC The sequences given in ABB80231-34 represent vp1 proteins derived from
 CC various adeno-associated virus (AAV) serotypes. These sequences were used
 CC in the scope of the invention for comparison with the cap protein derived
 CC from AAV serotype 9. The AAV capsid comprises three proteins vp1, vp2 and
 CC vp3, which are alternative splice variants. The AAV or the nucleic acid
 CC molecule is useful for preparing a medicament for delivering a transgene
 CC to a cell
 XX
 SQ Sequence 735 AA;
 Query Match 100.0%; Score 3994; DB 6; Length 735;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAADGYLPDWLEDTLSEGIROWWKLKPGPPPPKPAERHKDPSRLGLVLPVGYKLGPNGLD 60
 Db 1 MAADGYLPDWLEDTLSEGIROWWKLKPGPPPPKPAERHKDPSRLGLVLPVGYKLGPNGLD 60

QY 61 KGEVNEADAAALEHDKAYDROLDSGDNPLKYNHADADEFQERLKBEDTSFGNGLGRAVFQ 120
DB 61 KGEVNEADAAALEHDKAYDROLDSGDNPLKYNHADADEFQERLKBEDTSFGNGLGRAVFQ 120
QY 121 AKKRVLEPLGLVEEPVKTAPGKKRPVEHSPVPEPDSSTGKGAGQOPARKRLNFGQGDAD 180
DB 121 AKKRVLEPLGLVEEPVKTAPGKKRPVEHSPVPEPDSSTGKGAGQOPARKRLNFGQGDAD 180
QY 181 SVPDPQLGPPAAPASGLGTNTMATSGAPMADNNEGADGVNSGNWHCDSTWMDRVI 240
DB 181 SVPDPQLGPPAAPASGLGTNTMATSGAPMADNNEGADGVNSGNWHCDSTWMDRVI 240
QY 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHVFGYSTPGYDFNRFHCFSPRDMQRLI 300
DB 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHVFGYSTPGYDFNRFHCFSPRDMQRLI 300
QY 301 NNNMGFRPKRLNFKLNFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPVYLSAHQ 360
DB 301 NNNMGFRPKRLNFKLNFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPVYLSAHQ 360
QY 361 CLPPFPADVFMVPOQYGLTLNNGSQAAGRSSFYCYLEYFPQMLRTGNFTFSTFEDVVF 420
DB 361 CLPPFPADVFMVPOQYGLTLNNGSQAAGRSSFYCYLEYFPQMLRTGNFTFSTFEDVVF 420
QY 421 HSSYAHQSOLDRLMNPILDOYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDSRNWLP 480
DB 421 HSSYAHQSOLDRLMNPILDOYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDSRNWLP 480
QY 481 PCYRQORVSKTSADNNSEYSGTATKYHLNGRDSLNVNPGPAMASHKDEEKFPPQSGVL 540
DB 481 PCYRQORVSKTSADNNSEYSGTATKYHLNGRDSLNVNPGPAMASHKDEEKFPPQSGVL 540
QY 541 IFGKQSEKTNVDIEKVMITDEEIRTTNPVATEQYGSVSTNLQRNRQAATADVNTQGV 600
DB 541 IFGKQSEKTNVDIEKVMITDEEIRTTNPVATEQYGSVSTNLQRNRQAATADVNTQGV 600
QY 601 LFGMWQDRDQVYLOQPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
DB 601 LFGMWQDRDQVYLOQPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660

RESULT 6
ABR82436
ID ABR82436 standard; protein; 735 AA.
XX
AC ABR82436;
XX
DT 06-NOV-2003 (first entry)
XX
DE Adeno-associated virus 2 (AAV-2) CAP VP1 polypeptide.
XX
KW Adenovirus; Rep78; Rep68; gene therapy; AAV-2; CAP; VP1.
XX
OS Adeno associated virus.
XX
PN WO2003061582-A2.
XX
PD 31-JUL-2003.
XX
PF 17-JAN-2003; 2003WO-US001624.
XX
PR 18-JAN-2002; 2002US-0349532P.
XX
PA (UYDU-) UNIV DUKE.

XX Li C, Zhang X;
XX WPI; 2003-627412/59.
XX
XX New recombinant adenovirus comprising an adenovirus that encodes one or more AAV REP78/68 polypeptides that are inducibly expressed, useful as a vector for gene therapy.
XX
XX Claim 25; Page 112-115; 122pp; English.
XX
XX The invention relates to a recombinant adenovirus comprising an adenovirus that encodes one or more AAV REP78/68 polypeptides that are inducibly expressed. A complete virus-mediated system for recombinant AAV production is provided which comprises: (a) a first recombinant adenovirus encoding one or more AAV REP78/68 polypeptides and one or more viral capsid polypeptides; (b) a second recombinant adenovirus comprising a gene of interest and AAV inverted terminal repeats that flank the gene of interest; (c) viral helper functions; and (d) a host cell comprising the first recombinant adenovirus, the second recombinant adenovirus, and the viral helper functions. The recombinant AAV is useful as a vector for gene therapy. The present sequence represents an AAV-2 CAP VP1 polypeptide
XX
XX Sequence 735 AA;
XX
Query Match 100.0%; Score 3994; DB 6; Length 735;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAADGYLDWLEDDTJSEGIQWKKLPKPPPPKPAERHKDDSRGLVLPYKYLGPPNGLD 60
DB 1 MAADGYLDWLEDDTJSEGIQWKKLPKPPPPKPAERHKDDSRGLVLPYKYLGPPNGLD 60
QY 61 KGEVNEADAAALEHDKAYDROLDSGDNPLKYNHADADEFQERLKBEDTSFGNGLGRAVFQ 120
DB 61 KGEVNEADAAALEHDKAYDROLDSGDNPLKYNHADADEFQERLKBEDTSFGNGLGRAVFQ 120
QY 121 AKKRVLEPLGLVEEPVKTAPGKKRPVEHSPVPEPDSSTGKGAGQOPARKRLNFGQGDAD 180
DB 121 AKKRVLEPLGLVEEPVKTAPGKKRPVEHSPVPEPDSSTGKGAGQOPARKRLNFGQGDAD 180
QY 181 SVPDPQLGPPAAPASGLGTNTMATSGAPMADNNEGADGVNSGNWHCDSTWMDRVI 240
DB 181 SVPDPQLGPPAAPASGLGTNTMATSGAPMADNNEGADGVNSGNWHCDSTWMDRVI 240
QY 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHVFGYSTPGYDFNRFHCFSPRDMQRLI 300
DB 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHVFGYSTPGYDFNRFHCFSPRDMQRLI 300
QY 301 NNNMGFRPKRLNFKLNFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPVYLSAHQ 360
DB 301 NNNMGFRPKRLNFKLNFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPVYLSAHQ 360
QY 361 CLPPFPADVFMVPOQYGLTLNNGSQAAGRSSFYCYLEYFPQMLRTGNFTFSTFEDVVF 420
DB 361 CLPPFPADVFMVPOQYGLTLNNGSQAAGRSSFYCYLEYFPQMLRTGNFTFSTFEDVVF 420
QY 421 HSSYAHQSOLDRLMNPILDOYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDSRNWLP 480
DB 421 HSSYAHQSOLDRLMNPILDOYLYLSRTNTPSGTTTQSRLOFSQAGASDIRDSRNWLP 480
QY 481 PCYRQORVSKTSADNNSEYSGTATKYHLNGRDSLNVNPGPAMASHKDEEKFPPQSGVL 540
DB 481 PCYRQORVSKTSADNNSEYSGTATKYHLNGRDSLNVNPGPAMASHKDEEKFPPQSGVL 540
QY 541 IFGKQSEKTNVDIEKVMITDEEIRTTNPVATEQYGSVSTNLQRNRQAATADVNTQGV 600
DB 541 IFGKQSEKTNVDIEKVMITDEEIRTTNPVATEQYGSVSTNLQRNRQAATADVNTQGV 600
QY 601 LFGMWQDRDQVYLOQPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660
DB 601 LFGMWQDRDQVYLOQPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPSTT 660

QY 661 FSAAKFASFTQYSTGQSVIEIWELOKENS KRWNPBIQYTSNYSKVNVDFVTDTNGVY 720
 Db 661 FSAAKFASFTQYSTGQSVIEIWELOKENS KRWNPBIQYTSNYSKVNVDFVTDTNGVY 720
 QY 721 SEPRPIGTRYLTRNL 735
 Db 721 SEPRPIGTRYLTRNL 735

RESULT 7
 ABR62761
 ID ABR62761 standard; protein; 735 AA.
 AC ABR62761;
 XX 06-NOV-2003 (first entry)
 DT Adeno associated virus 2 capsid protein vpl.
 XX AAV; AAV2; capsid; vector; gene therapy; antisense therapy; vaccine.
 KW Adeno associated virus.
 OS
 XX Key Location/Qualifiers
 FH Region 1. .184
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 24. .42
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 25. .28
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 81. .85
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 133. .165
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 134. .165
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 137. .143
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 146. .152
 FT /note= "hypervariable region, specifically referred to in
 FT Claim 4"
 FT Region 154. .156
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 171. .173
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 182. .187
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 185. .198
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 185. .198
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 194. .208
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 199. .259
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 260. .273
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"

FT Region 261. .274
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 262. .274
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 262. .264
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 263. .266
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 274. .446
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 381. .383
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 383. .385
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 413. .417
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 447. .477
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 449. .478
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 450. .474
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 451. .475
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 490. .495
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 491. .496
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 494. .525
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 495. .602
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 500. .504
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 501. .505
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 514. .522
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 533. .554
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 534. .571
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 534. .555
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 581. .601
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 581. .594
 FT /note= "hypervariable region specifically referred to in
 FT Claim 4"
 FT Region 583. .596

XX 12-NOV-2002; 2002EP-00257826.
 XX 13-NOV-2001; 2001US-0350607P.
 PR 17-DEC-2001; 2001US-0341117P.
 PR 01-MAY-2002; 2002US-0377066P.
 PR 05-JUN-2002; 2002US-0386675P.
 XX (UYPE-) UNIV PENNSYLVANIA.
 XX Gao G, Wilson JM, Alvira M;
 XX WPI; 2003-450984/43.
 XX Detecting adeno-associated virus sequences in a sample, useful for e.g.
 .PT preventing or treating hyperproliferative or autoimmune diseases,
 .PT comprises subjecting a sample having a DNA to amplification via
 .PT polymerase chain reaction.
 XX Disclosure; SEQ ID NO 70; 419pp; English.
 .PS
 .XX
 .XX The invention relates to a novel method for detecting adeno-associated
 CC virus (AAV) sequences in a sample, which comprises subjecting a sample
 CC containing a DNA to amplification via a polymerase chain reaction (PCR).
 CC The AAV sequence have the following activities: cytostatic,
 CC antipsoriatic, antirheumatic, antiarthritic, neuroprotective,
 CC antidiabetic, antithyroid, dermatological, and antiinflammatory. The AAV
 CC sequence can be used in gene therapy or as part of a vaccine to treat
 CC disorders. The method is useful in detecting and/or identifying AAV
 CC sequences and isolating novel sequences that are identified. The
 CC sequences may be used e.g. for preventing or treating hyperproliferative
 CC conditions such as cancers and psoriasis, and other autoimmune diseases
 CC like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune
 CC thyroiditis, scleroderma or Crohn's disease. This sequence represents an
 CC AAV related protein sequence of the invention.
 XX
 XX Sequence 735 AA;
 SQ

Query Match 100.0%; Score 3994; DB 7; Length 735;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 MAAGYLDDWLEDTLSEGIROWKLPKPPPPKPAERHKODSRLGLVLPYKYLGPFGNGLD 60
 1 MAAGYLDDWLEDTLSEGIROWKLPKPPPPKPAERHKODSRLGLVLPYKYLGPFGNGLD 60
 61 KGEFVNEADAALHDXAYDRQLSDGNPVLKYNHADAERQERLKEDTSFGNIGRAVFQ 120
 61 KGEFVNEADAALHDXAYDRQLSDGNPVLKYNHADAERQERLKEDTSFGNIGRAVFQ 120
 121 AKKRVLPGLVPEPVKTAPEGKRPVSHSPVEPDSSSGTGKAGQAPARKLNFGQTGDAD 180
 121 AKKRVLPGLVPEPVKTAPEGKRPVSHSPVEPDSSSGTGKAGQAPARKLNFGQTGDAD 180
 181 SVDPDQPLGPPAPSPGLTNTWATGSGAPADNNEGADGVNNGWCHDSTWMDRVI 240
 181 SVDPDQPLGPPAPSPGLTNTWATGSGAPADNNEGADGVNNGWCHDSTWMDRVI 240
 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPWCYFDFNRFCHFSPRDWORLI 300
 241 TTSTRTWALPTYNHLYKQISSQSGASNDNHFGYSTPWCYFDFNRFCHFSPRDWORLI 300
 301 NNNWGPFPKLNPKLFNIQVEVTQNDGTTIANNLTSTVQVFTDSEYQLPYVLGSAHQG 360
 301 NNNWGPFPKLNPKLFNIQVEVTQNDGTTIANNLTSTVQVFTDSEYQLPYVLGSAHQG 360
 361 CLPPFPADVFWPQYGYLTILNNGSQAVGRSSFFCLEYFSPQMLRTGNFTFSFTFEDVDF 420
 361 CLPPFPADVFWPQYGYLTILNNGSQAVGRSSFFCLEYFSPQMLRTGNFTFSFTFEDVDF 420
 421 HSSYAHQSQSLDRMLNPLIDLYLYLGRNTNTPSGTTQSRLOQFSQAGASDIRQSRNLPFG 480
 421 HSSYAHQSQSLDRMLNPLIDLYLYLGRNTNTPSGTTQSRLOQFSQAGASDIRQSRNLPFG 480

QY 481 PCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNPGPAMASHKDDDEKFPQSGVL 540
 DB 481 PCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNPGPAMASHKDDDEKFPQSGVL 540
 QY 541 IFGQKQSEKTNVDIEKVMITDEBEIRTNPNVATEQYGSVSTNLQGRNRQAATADVNTQGV 600
 DB 541 IFGQKQSEKTNVDIEKVMITDEBEIRTNPNVATEQYGSVSTNLQGRNRQAATADVNTQGV 600
 QY 601 LPGMVWODRDVYLGQPIWAKIPHTDGHFHSPLMGFGGLKHPPOILLIKNTVPANPSTT 660
 DB 601 LPGMVWODRDVYLGQPIWAKIPHTDGHFHSPLMGFGGLKHPPOILLIKNTVPANPSTT 660
 QY 661 FSAAKFASFTQYSTQGVSVIEIWEIWKQENSKENWPEIQYTSNYSKSVNVDFTVDTNGVY 720
 DB 661 FSAAKFASFTQYSTQGVSVIEIWEIWKQENSKENWPEIQYTSNYSKSVNVDFTVDTNGVY 720
 QY 721 SEPRPIGTRYLTRNL 735
 DB 721 SEPRPIGTRYLTRNL 735

RESULT 9
 AAB50326
 ID AAB50326 standard; protein; 734 AA.
 XX AAB50326;
 AC AAB50326;
 DT 09-MAR-2001 (first entry)
 XX Adeno-associated virus capsid protein sequence.
 DE Adeno-associated virus; AAV; capsid; virus binding inhibition;
 XX Adeno-associated virus; AAV; capsid; virus binding inhibition;
 KW competitive inhibitor.
 KW Mastadenovirus.
 OS Synthetic.
 XX WO2000073316-A2.
 PN 07-DEC-2000.
 XX 26-MAY-2000; 2000WO-US014466.
 PF 28-MAY-1999; 99US-00321589.
 PR (CELL-) CELL GENESYS INC.
 PA Patel S, McArthur J;
 XX WPI; 2001-061507/07.
 XX New polypeptide portion of a virus protein binding to an antibody
 PT specific for the virus useful for inhibiting binding of the virus to a
 PT cell or for binding host antibody to provide a transient tolerant or non-
 PT responsive state.
 XX Disclosure; Fig 2; 33pp; English.
 PS The present sequence is given in a specification relating to polypeptide
 CC portions of a virus protein or its derivative, that bind to an antibody
 CC specific for the virus or inhibit binding of the virus to a cell. The
 CC polypeptides are used in inhibiting the binding of viruses to cells of a
 CC host. Oligopeptides that inhibit binding of virus to receptor can be used
 CC as competitive inhibitors to release bound virus in an adsorption-type
 CC assay, and if an antibody was used as an immunoadsorbent, the
 CC oligopeptide could be used to elute bound virus from a solid support
 CC which virus antibody is immobilised. These oligopeptides may further be
 CC used to bind to host antibody to provide a transient tolerant or non-
 CC responsive state
 XX Sequence 734 AA;
 SQ

XX	OS	Unidentified.
XX	PN	EPI310571-A2.
XX	PD	14-MAY-2003.
XX	PF	12-NOV-2002; 2002EP-00257826.
XX	PR	13-NOV-2001; 2001US-0350607P.
XX	PR	17-DEC-2001; 2001US-0341117P.
XX	PR	01-MAY-2002; 2002US-0377066P.
XX	PR	05-JUN-2002; 2002US-0386675P.
XX	PA	(UYPE-) UNIV PENNSYLVANIA.
XX	PI	Gao G, Wilson JM, Alvira M;
XX	DR	WPI; 2003-450984/43.
XX	PT	Detecting adeno-associated virus sequences in a sample, useful for e.g. preventing or treating hyperproliferative or autoimmune diseases,
XX	PT	comprises subjecting a sample having a DNA to amplification via
XX	PT	polymerase chain reaction.
XX	PS	Claim 24; SEQ ID NO 67; 419pp; English.
XX	CC	The invention relates to a novel method for detecting adeno-associated
XX	CC	virus (AAV) sequences in a sample, which comprises subjecting a sample
XX	CC	containing a DNA to amplification via a polymerase chain reaction (PCR).
XX	CC	The AAV sequence have the following activities: cytostatic,
XX	CC	antiproliferative, antirheumatic, antiarthritic, neuroprotective,
XX	CC	antidiabetic, antithyroid, dermatological, and antiinflammatory. The AAV
XX	CC	sequence can be used in gene therapy or as part of a vaccine to treat
XX	CC	disorders. The method is useful in detecting and/or identifying AAV
XX	CC	sequences and isolating novel sequences that are identified. The
XX	CC	sequences may be used e.g. for preventing or treating hyperproliferative
XX	CC	conditions such as cancers and psoriasis, and other autoimmune diseases
XX	CC	like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune
XX	CC	thyroiditis, scleroderma or Crohn's disease. This sequence represents an
XX	CC	AAV related protein sequence of the invention.
XX	SQ	Sequence 735 AA;
Query Match		
Best Local Similarity 90.7%; Score 3623; DB 7; Length 735;		
Matches 655; Conservative 38; Mismatches 42; Indels 0; Gaps 0;		
Qy	1	MAADGYLPDWMLEDTLSEGIQRQWKXKPGPPPKPAERHKDDSRGLVPGYKYLGPPNGLD 60
Db	1	MAADGYLPDWMLEDTLSEGIQRQWKXKPGPPPKPNQHRDSDRGLVPGYKYLGPPNGLD 60
Qy	61	KGEVNEADAALAEHDKAYDRLDSGDNPYLKYNHADAEFQBRLEDTSFGNLGRAVFQ 120
Db	61	KGEVNEADAALAEHDKAYDQLXQGDNPYLKYNHADAEFQBRLEQDTSFGNLGRAVFQ 120
Qy	121	AKRVLRPLGLVEEPPVKTPACKRKRPVHSVPEPDSSTGGTGAGQQAPARKLNFGQTGDAD 180
Db	121	AKRVLRPLGLVEEAVKTPACKRKPIESPAEDPSSGIGKGQQAPAKRLNFGQTGDIYE 180
Qy	181	SVPDPQLGPQPAAPSGSLGTINTMATGSAPMADNNENAGDVGNSSGNWHCDSTWMDRVI 240
Db	181	SVPDPQIPEPAAPSGVSNMASGGCAPMADNNENAGDVGNSSGNWHCDSTWMDRVI 240
Qy	241	TTSTRTRTWTALPTYNNHLKYQISSQGASNDNHFYGYSTPWGYPDFNRHFCHFSFRDWORLI 300
Db	241	TTSTRTRTWTALPTYNNRLKYQISSEGATNDNHIFYGYSTPWGYPDFNRHFCHFSFRDWORLI 300
Qy	301	NNNWGFRRPKRLFKNLFNIQVKEVTQNDCGTTTTIANLLTSTVOVFTDSEYOLPYVLGSAHQG 360
Db	301	NNNWGFRRPKRLFKNLFNIQVKEVTQNDCGTTTTIANLLTSTVOVFTDSEYQLPVVLGSAHQG 360
Qy	361	CLPPFPADVFMVPOGYILTLNNGSOAVGRSSFYCLEYFFPSQMLRTGCNNFTSFYTDEDVFP 420

ADP76568

XX

XX

DE
XX

KW aden

KW antiarthritic; neuroprotective; antidiabetic; antithyroid;
KW dermatologic; antiinflammatory; gene therapy; vaccine;
KW hyperproliferative; cancer; psoriasis; autoimmune disease;
KW rheumatoid arthritis; multiple sclerosis; diabetes;
KW autoimmune thyroiditis; scleroderma; Crohn's disease.

361 CLPFPADVFMIPQYGLYLNNQSVAGRSSFCLEFFSQQMLRTGNNTFTTTFEDVPF 420
 421 HSSYAHQSQSLDRMLNPLIDQYLYLRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
 421 HSSYAHQSQSLDRMLNPLIDQYLYLRTNTPSGTTTQSRLOFSQAGPSSMAQAQKNWLP 480
 481 PCYRQQRVSKTSADNNNSYSWTGATKYHLNGRDSLVNPGPAMASHKODEEKFPPQSGVL 540
 481 PSYRQQRMSKTANDNNNSSEFAWTAATKYLNGRNSLVNPGPAMASHKODEEKFPMHGNL 540
 541 IFGKQSEKTNVDIEKVMITDEEIRITNPVATEQYGSVSTNLQRNQAAATADVNTQGV 600
 541 IFGKQGTGTNTVDIESVLITDEEIRITNPVATEQYGVATNHQSNNTASVGSVDSQGI 600
 601 LFGMWQDRDVLQGIPIWAKI PHTDGHFHPSPLMGGFGLKHPHPQIILKNTVPANPSTT 660
 601 LFGMWQDRDVLQGIPIWAKI PHTDGHFHPSPLMGGFGLKHPHPQIILKNTVPANPSTT 660
 661 FSAAKFASFIQYSTQGVSVIEWELQKENSKRWNPEIQYTSNKNYSVNVFTVDTNGVY 720
 661 FPGKFAFIQYSTQGVSVIEWELQKENSKRWNPEIQYTSNKNYSVNVFTVDTNGVY 720
 721 SEPRPIGTRYLTRNL 735
 721 SEPRPIGTRYLTRNL 735

RESULT 11
 AD76570
 ID AD76570 standard; protein; 735 AA.
 AC AD76570;
 XX
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Adeno-associated virus (AAV) related protein, SEQ ID No 69.
 XX
 KW adeno-associated virus; AAV; cytostatic; antipariatic; antirheumatic;
 KW antarthritic; neuroprotective; antidiabetic; antithyroid;
 KW dermatological; antinflammatory; gene therapy; vaccine;
 KW hyperproliferative; cancer; psoriasis; autoimmune disease;
 KW rheumatoid arthritis; multiple sclerosis; diabetes;
 KW autoimmune thyroiditis; scleroderma; Crohn's disease.
 XX
 OS Unidentified.
 XX
 PN EP1310571-A2.
 XX
 PD 14-MAY-2003.
 XX
 PF 12-NOV-2002; 2002EP-00257826.
 XX
 PR 13-NOV-2001; 2001US-0350607P.
 PR 17-DEC-2001; 2001US-0341117P.
 PR 01-MAY-2002; 2002US-0377066P.
 PR 05-JUN-2002; 2002US-0386675P.
 XX
 PA (UNP) UNIV PENNSYLVANIA.
 XX
 PI Gao G, Wilson JM, Alvira M;
 XX
 DR WPI; 2003-450984/43.
 XX
 XX
 PT Detecting adeno-associated virus sequences in a sample, useful for e.g.
 PT preventing or treating hyperproliferative or autoimmune diseases,
 FT comprises subjecting a sample having a DNA to amplification via
 PT polymerase chain reaction.
 XX
 XX
 PS Claim 24; SEQ ID NO 69; 419pp; English.
 XX
 CC The invention relates to a novel method for detecting adeno-associated
 CC virus (AAV) sequences in a sample, which comprises subjecting a sample

CC containing a DNA to amplification via a polymerase chain reaction (PCR).
 CC The AAV sequence have the following activities: cytostatic,
 CC antipariatic, antirheumatic, antarthritic, neuroprotective,
 CC antidiabetic, antinflammatory, dermatological, and antinflammatory. The AAV
 CC sequence can be used in gene therapy or as part of a vaccine to treat
 CC disorders. The method is useful in detecting and/or identifying AAV
 CC sequences and isolating novel sequences that are identified. The
 CC sequences may be used e.g. for preventing or treating hyperproliferative
 CC conditions such as cancers and psoriasis, and other autoimmune diseases
 CC like rheumatoid arthritis, multiple sclerosis, diabetes, autoimmune
 CC thyroiditis, scleroderma or Crohn's disease. This sequence represents an
 CC AAV related protein sequence of the invention.
 XX
 SQ Sequence 735 AA;

Query Match 90.7%; Score 3622; DB 7; Length 735;
 Best Local Similarity 89.0%; Pred. No. 9.4e-289;
 Matches 654; Conservative 39; Mismatches 42; Indels 0; Gaps 0;
 QY 1 MAADGYLPDLWLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPYKYLGPENGLD 60
 DB 1 MAADGYLPDLWLEDTLSEGIQWMLKPGPPPPKPAERHKDDSRGLVLPYKYLGPENGLD 60
 QY 61 KGEVNEADAALAHDKAYDRLDSDGNPYLYKNHADAFOERLKEDETSFGNGLGRAVQ 120
 DB 61 KGEVNEADAALAHDKAYDRLDSDGNPYLYKNHADAFOERLKEDETSFGNGLGRAVQ 120
 QY 121 AKKRVLEPLGLVEEVPVKTAPEKRPVVEHSPVPSDSSGTGKAGQQAPARLNFQGTAD 180
 DB 121 AKKRVLEPLGLVEEVPVKTAPEKRPVVEHSPVPSDSSGTGKAGQQAPARLNFQGTAD 180
 QY 181 SVDPDQPLQGPAPPSGLGTNTMATSGAPMADNNEGADGVNNSGNHCHDSTWMDRVI 240
 DB 181 SVDPDQPLQGPAPPSGLGTNTMATSGAPMADNNEGADGVNNSGNHCHDSTWMDRVI 240
 QY 241 TTSTRTWALPTNNHLYKQISSQSGASNDNHVFGYSTPMGYDFNRFHCHESPRDWR 300
 DB 241 TTSTRTWALPTNNHLYKQISSQSGASNDNHVFGYSTPMGYDFNRFHCHESPRDWR 300
 QY 301 NNNWGRPRKRLNPKLFNIQVKEVTQNDGTTTANNLTSTVQVFTDSEYQLPYVLSAHQ 360
 DB 301 NNNWGRPRKRLNPKLFNIQVKEVTQNDGTTTANNLTSTVQVFTDSEYQLPYVLSAHQ 360
 QY 361 CLPFPADVFMIPQYGLYLNNQSVAGRSSFCLEFFSQQMLRTGNNTFTTTFEDVPF 420
 DB 361 CLPFPADVFMIPQYGLYLNNQSVAGRSSFCLEFFSQQMLRTGNNTFTTTFEDVPF 420
 QY 421 HSSYAHQSQSLDRMLNPLIDQYLYLRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
 DB 421 HSSYAHQSQSLDRMLNPLIDQYLYLRTNTPSGTTTQSRLOFSQAGASDIRDQSRNWLPG 480
 QY 481 PCYRQQRVSKTSADNNNSYSWTGATKYHLNGRDSLVNPGPAMASHKODEEKFPPQSGVL 540
 DB 481 PCYRQQRVSKTSADNNNSYSWTGATKYHLNGRDSLVNPGPAMASHKODEEKFPPQSGVL 540
 QY 541 IFGKQSEKTNVDIEKVMITDEEIRITNPVATEQYGSVSTNLQRNQAAATADVNTQGV 600
 DB 541 IFGKQGTGTNTVDIESVLITDEEIRITNPVATEQYGVATNHQSNNTASVGSVDSQGI 600
 QY 601 LFGMWQDRDVLQGIPIWAKI PHTDGHFHPSPLMGGFGLKHPHPQIILKNTVPANPSTT 660
 DB 601 LFGMWQDRDVLQGIPIWAKI PHTDGHFHPSPLMGGFGLKHPHPQIILKNTVPANPSTT 660
 QY 661 FSAAKFASFIQYSTQGVSVIEWELQKENSKRWNPEIQYTSNKNYSVNVFTVDTNGVY 720
 DB 661 FPGKFAFIQYSTQGVSVIEWELQKENSKRWNPEIQYTSNKNYSVNVFTVDTNGVY 720
 QY 721 SEPRPIGTRYLTRNL 735
 DB 721 SEPRPIGTRYLTRNL 735

QY 720 YSEPRPIGTRYLNRNL 735
|||
Db 721 YSEPRPIGTRYLNRNL 736

Search completed: May 19, 2005, 13:30:02
Job time : 169 secs